

UNIVERSIDADE FEDERAL DO PARANÁ

MATHEUS ASSIS DOMINGUES

DESENVOLVIMENTO DE UM SENSOR VIRTUAL
PARA CONTROLE DA RESISTÊNCIA À TRAÇÃO DO
PAPEL EM UMA PLANTA DE POLPA CTMP

CURITIBA-PR

2019

MATHEUS ASSIS DOMINGUES

DESENVOLVIMENTO DE UM SENSOR VIRTUAL
PARA CONTROLE DA RESISTÊNCIA À TRAÇÃO DO
PAPEL EM UMA PLANTA DE POLPA CTMP

Dissertação apresentado ao curso de Pós-Graduação em Engenharia e Ciência dos Materiais, Setor de tecnologia, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Engenharia e Ciência dos Materiais.

Orientador: Prof. Dr. Carlos Itsuo Yamamoto

Co-orientador: Prof. Dr. Osvaldo Vieira

CURITIBA-PR

2019

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

D671d Domingues, Matheus Assis
Desenvolvimento de um sensor virtual para controle da resistência à tração do papel em uma planta de polpa CTMP [recurso eletrônico] /Matheus Assis Domingues. – Curitiba, 2019.

Dissertação – Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia e Ciência dos Materiais, 2019.

Orientador: Carlos Itsuo Yamamoto. Coorientador: Osvaldo Vieira.

1. Celulose. 2. Papel. 3. Redes neurais (Computação). I. Universidade Federal do Paraná. II. Yamamoto, Carlos Itsuo. III. Vieira, Osvaldo. IV. Título.

CDD: 572.56682

Bibliotecária: Vanusa Maciel CRB- 9/1928



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA E
CIÊNCIA DOS MATERIAIS - 40001016033P9

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA E CIÊNCIA DOS MATERIAIS da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **MATHEUS ASSIS DOMINGUES** intitulada: **DESENVOLVIMENTO DE UM SENSOR VIRTUAL PARA CONTROLE DA RESISTÊNCIA À TRAÇÃO DO PAPEL EM UMA PLANTA DE POLPA CTMP**, sob orientação do Prof. Dr. CARLOS ITSUO YAMAMOTO, que após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

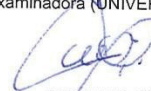
A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 23 de Agosto de 2019.



CARLOS ITSUO YAMAMOTO

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)



JOSE LUIS FABRIS

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ)



MARCELO KAMINSKI LENZI

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Agradeço a disponibilidade do professor Dr. Carlos Yamamoto pela orientação durante o mestrado, contribuindo para a realização do trabalho e por toda confiança.

A Klabin S.A, pela possibilidade de execução desse trabalho, disponibilizando sua estrutura e por todo suporte técnico dos funcionários, Obrigado em especial ao meu coorientador e engenheiro da Klabin Dr. Osvaldo Vieira, que sempre teve disponibilidade para ajudar.

Ao Programa de Pós-Graduação em Engenharia dos Materiais (PIPE), a UFPR e ao Laboratório LACAUT pela infraestrutura disponibilizada, em especial os funcionários Renato, Rubia e Cris por toda ajuda durante esta jornada. Agradeço igualmente a Coordenação de Aperfeiçoamento de Nível Pessoal (CAPES) pelo financiamento da presente pesquisa.

Minha mãe Neuza, meu pai Jânio e a minha namorada Ana Paula por estarem presentes em grande parte da minha caminhada. A minha família que mesmo distante foi minha força nos momentos difíceis através de encorajamento, ajuda e preocupação durante este período.

A todos que direta ou indiretamente colaboraram nesta jornada, em especial a Sarah, Daniela, Priscila, Heloisa, Filipe, Gabriela e Lúrima, que sempre me deram apoio e incentivo para continuar este projeto e por sempre me escutar e dar conselhos nas horas mais críticas.

RESUMO

Na indústria de celulose e papel, no Brasil, as principais fontes de matéria prima para a produção industrial são provenientes das espécies arbóreas Eucaliptos e Pinus. A partir do processo de polpação é produzido a polpa de celulose para a produção do papel, um dos métodos de polpação é o *chemi-thermomechanical pulp* e esse processo possui grande influência nas propriedades do produto final. Para atender as demandas do mercado, é necessário um grande controle sobre o processo de produção, e um dos instrumentos utilizados para esse controle são os sensores físicos, que possuem um alto custo para as empresas. Uma alternativa para esses sensores físicos são os sensores virtuais, que se utiliza de modelos matemáticos para predição da variável de interesse, e possuem baixo custo de desenvolvimento, instalação e manutenção além de uma alta precisão. Este trabalho tem como objetivo desenvolver um sensor virtual para a predição da resistência à tração do papel-cartão produzido pela empresa KLABIN. Utilizou-se de modelagem estatística para o desenvolvimento de um modelo por regressão linear multivariado e um modelo por regressão não linear por mínimos quadrados parciais (PLS). Desenvolveu-se um modelo baseado em redes neurais artificiais (RNA) a fim de comparar o desempenho do mesmo com os modelos estatísticos e selecionar o melhor modelo para o desenvolvimento do sensor virtual, os dados para o desenvolvimento do modelo foram divididos em 60% para treinamento, 20% para validação e 20% para teste. Os modelos desenvolvidos apresentaram a princípio um baixo desempenho na previsão da variável de interesse, indicando problemas no banco de dados. *Principal Componente Analysis* foi aplicado no conjunto de dados a fim de resolver o problema, os modelos desenvolvidos após a utilização de PCA apresentaram novamente um resultado não satisfatório, tendo desempenho inferior aos modelos desenvolvidos anteriormente. A partir do grande poder de generalização das redes neurais e do seu alto grau de não-linearidade, removeu-se *outliers* do conjunto de dados, formando um novo *dataset* para o desenvolvimento dos modelos. Os modelos apresentaram desempenho superior aos modelos anteriores, o modelo baseado em redes neurais foi o que apresentou maior capacidade de predizer a variável de interesse com $R = 0,89$.

Palavras-chave: Celulose, Papel, Modelagem Matemática, Redes Neurais Artificiais, Sensores Virtuais.

ABSTRACT

In the Brazilian pulp and paper industry, the main sources of raw material for industrial production come from the tree species *Eucalyptus* and *Pinus*. The pulp produced for paper production comes from the pulping process, one of the pulping methods is the chemi-thermomechanical pulp and this process has great influence on the properties of the final product. To achieve the demands of the Market, a tight control of the production process is necessary, and one of the instruments used to control are physical sensors, which have high costs to the companies. An alternative for these physical sensors are soft sensors, which use mathematical models to predict the variable of interest, with low development, installation and maintenance costs in addition to high precision. This work aims to develop a soft sensor for the prediction of tensile strength of paperboard produced by KLABIN. Statistical modeling was used to develop a multivariate-linear regression model and a non-linear regression model based on partial least squares (PLS). A model based on neural network (NN) was developed in order to compare the performance of the NN model with the statistical models and selecting the best model for the development of the soft sensor, the dataset used to develop the NN was divided into 60% for training, 20% for validation and 20% for testing. The models developed presented at first a poor performance in the prediction of the variable of interest, indicating problems in the dataset. Principal Component Analysis was applied to the dataset in order to solve the problem, the models developed after the use of PCA again presented an unsatisfactory result, having inferior performance compared to the previously developed models. Due to the capacity of neural network to generalize and its high nonlinearity, NN was used as a method to identify outliers on the dataset, forming a new one that was used to develop new models. The models performed better than the previous models developed and the model based on NN presented the highest capacity to predict the variable of interest with $R = 0,89$.

Keywords: Pulp, Paper, Mathematical Modeling, Neural Network, Soft Sensor

LISTA DE FIGURAS

FIGURA 1 - TESTE DE RESISTÊNCIA À TRAÇÃO APLICADO SOBRE O PAPEL	22
FIGURA 2 - DINAMÔMETRO UTILIZADO PARA CALCULAR A RESISTÊNCIA À TRAÇÃO	22
FIGURA 3 - FLUXOGRAMA DA PLANTA DE PRODUÇÃO CTMP.....	24
FIGURA 4 - CLUSTERING REPRESENTADO POR MEIO DO DENDOGRAMA	32
FIGURA 5 - REPRESENTAÇÃO DE UM NEURÔNIO BIOLÓGICO.....	36
FIGURA 6 - REPRESENTAÇÃO DE UM NEURÔNIO ARTIFICIAL.....	37
FIGURA 7 – ARQUITETURA <i>FEEDFORWARD</i> COM UMA CAMADA OCULTA.....	41
FIGURA 8 - NEURÔNIO ARTIFICIAL COM UMA CAMADA COM ESTRUTURA RECORRENTE.....	42
FIGURA 9 - ESTRUTURA DO TREINAMENTO SUPERVISIONADO BACKPROPAGATION DE UMA REDE DE MÚLTIPLAS CAMADAS.....	43
FIGURA 10 - DIAGRAMA DE PRINCÍPIO DE SENSOR VIRTUAL.....	50
FIGURA 11 - ANÁLISE DE CORRELAÇÃO LINEAR PELO MÉTODO DE PEARSON.....	58
FIGURA 12 - ANÁLISE DE CORRELAÇÃO NÃO-LINEAR PELO MÉTODO DE SPEARMAN.....	59
FIGURA 13 - ESTRUTURA DA REDE NEURAL ARTIFICIAL DESENVOLVIDA.....	64
FIGURA 14 - ESTADO DO TREINAMENTO DA REDE NEURAL ARTIFICIAL.....	64

FIGURA 15 - REGRESSÃO DO MODELO BASEADO EM REDES NEURAIIS.....	65
FIGURA 16 - COMPARAÇÃO DOS RESULTADOS DA REDE E DOS RESULTADOS REAIS.....	66
FIGURA 17 - HISTOGRAMA DE ERRO DA REDE NEURAL ARTIFICIAL.....	67
FIGURA 18 - ESTRUTURA DA REDE NEURAL ARTIFICIAL (PCA)	70
FIGURA 19 - ESTADO DO TREINAMENTO DA REDE NEURAL ARTIFICIAL (PCA).....	70
FIGURA 20 - REGRESSÃO DO MODELO BASEADO EM REDES NEURAIIS (PCA).....	71
FIGURA 21 - COMPARAÇÃO ENTRE A RNA E OS RESULTADOS REAIS (PCA).....	72
FIGURA 22 - HISTOGRAMA DE ERRO DA REDE NEURAL ARTIFICIAL (PCA).....	73
FIGURA 23 - ESTRUTURA DA REDE NEURAL ARTIFICIAL (CONJUNTO DE DADOS RNA)	77
FIGURA 24 - ESTADO DO TREINAMENTO DA REDE NEURAL (CONJUNTO DE DADOS RNA)	77
FIGURA 25 - REGRESSÃO DO MODELO BASEADO EM REDES NEURAIIS (CONJUNTO DE DADOS RNA)	78
FIGURA 26 - COMPARAÇÃO ENTRE A RNA E OS RESULTADOS REAIS (CONJUNTO DE DADOS RNA)	79
FIGURA 27 - HISTOGRAMA DE ERRO DA REDE NEURAL ARTIFICIAL (CONJUNTO DE DADOS RNA)	80
FIGURA 28 - ANÁLISE DE CORRELAÇÃO NÃO-LINEAR PELO MÉTODO DE SPEARMANA NO NOVO <i>DATASET</i>	81

LISTA DE TABELAS

TABELA 1 - VISÃO SIMPLIFICADA DO CAMPO DE ANÁLISE MULTIVARIÁVEL.....	26
TABELA 2 - PRINCIPAIS FUNÇÕES DE ATIVAÇÃO PARCIALMENTE DIFERENCIÁVEIS.....	39
TABELA 3 - PRINCIPAIS FUNÇÕES DE ATIVAÇÃO TOTALMENTE DIFERENCIÁVEIS.....	40
TABELA 4 - ESTADO DA ARTE DA UTILIZAÇÃO DE REDES NEURAIS NA INDÚSTRIA DE CELULOSE E PAPEL.....	46
TABELA 5 - NÚMERO DE ELEMENTOS ENCONTRADO EM CADA CLUSTER.....	57
TABELA 6 - VARIÁVEIS SELECIONADAS PARA O DESENVOLVIMENTO DOS MODELOS.....	60
TABELA 7 - PARÂMETROS DO MODELO LINEAR MULTIVARIADO.....	61
TABELA 8 - ANÁLISE DO MODELO DE REGRESSÃO LINEAR MULTIPLA.....	62
TABELA 9 - PARÂMETROS DO MODELO PLS.....	62
TABELA 10 - ANÁLISE DO MODELO PLS.....	63
TABELA 11 – PORCENTAGEM EXPLICATIVA DE CADA VARIÁVEL.....	68
TABELA 12 – PARÂMETROS DO MODELO LINEAR MULTIVARIADO (PCA).....	68
TABELA 13 – ANÁLISE DO MODELO DE REGRESSÃO LINEAR .MULTIVARIADO (PCA).....	69
TABELA 14 – PARÂMETROS DO MODELO PLS (PCA)	69
TABELA 15 – PARÂMETROS DO MODELO LINEAR MULTIVARIADO (RNA).....	74

TABELA 16 – ANÁLISE DO MODELO DE REGRESSÃO LINEAR MULTIVARIADO (RNA).....	75
TABELA 17 – PARÂMETROS DO MODELO PLS (RNA).....	75
TABELA 18 – ANÁLISE DO MODELO PLS (RNA)	76
TABELA 19 – ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS DO PROCESSO.....	81

LISTA DE ABREVIATURAS OU SIGLAS

CAR	-	Comprimento de Auto Ruptura
CTMP	-	<i>Chemi-Thermomechanical Pulp</i>
ETP	-	<i>Effluent Total Phosphorus</i>
FT-NIR	-	<i>Fourrier Transform Near Infrared</i>
GUI	-	<i>Graphical User Interface</i>
HHM	-	<i>Hidden Markov Model</i>
LACAUT	-	Laboratório de Análises de Combustíveis Automotivos
MAG	-	Modelos Aditivos Generalizados
MLP	-	<i>Multilayer Perceptron</i>
NIPALS	-	<i>Nonlinear Iterative Partial Least Squares</i>
PCA	-	<i>Principal Components Analysis</i>
PCR	-	<i>Principal Components Regression</i>
PLS	-	<i>Partial Least Squares</i>
RBFNN	-	<i>Radial Basis Function Neural Network</i>
RMSE	-	<i>Root Mean Square Error</i>
RNA	-	Redes Neurais Artificiais

LISTA DE SÍMBOLOS

C_R	-	Carga de Ruptura
G	-	Gramatura
L	-	Largura
U	-	Potencial de ativação
d_i	-	Diferença Significativa
r_p	-	Correlação de Pearson
r_s	-	Correlação de Spearman
ρ	-	Correlação de Pearson entre os <i>ranks</i> das variáveis
θ	-	<i>Bias</i>

SUMÁRIO

1. INTRODUÇÃO	16
1.1 JUSTIFICATIVAS	17
1.2 OBJETIVO GERAL	18
1.3 OBJETIVOS ESPECÍFICOS	18
1.4 ORGANIZAÇÃO DA DISSERTAÇÃO	18
2. REVISÃO BIBLIOGRÁFICA	20
2.1 PROPRIEDADES DO PAPEL	20
2.2 PROCESSO DE PRODUÇÃO DE UMA PLANTA CTMP	23
2.3 MODELOS DE PREDIÇÃO	25
2.4 MODELOS DE PREDIÇÃO ESTATÍSTICOS	27
2.4.1 Análise de Variáveis	27
2.4.2 Regressão Linear Multivariável	28
2.4.3 Regressão por Mínimos Quadrados Parciais	29
2.4.4 Análise de <i>Cluster</i>	31
2.4.5 Análise dos Componentes Principais	33
2.5 REDES NEURAIS ARTIFICIAIS	36
2.5.1 Tipos de Arquiteturas, Treinamento de Redes Neurais e Validação	41
2.5.2 Modelos Baseados em Redes Neurais Aplicados na Indústria de Celulose e Papel	44
2.5.3 Uso de Redes Neurais Para a Detecção de <i>Outliers</i>	47
2.6 SENSORES VIRTUAIS	49
3. MATERIAIS E MÉTODOS	52
3.1. MÉTODOS	52
3.1.1. Identificação das Propriedades	52
3.1.2. Pré-Processamento de Dados	52
3.1.3. Análise Estatística de Variáveis	53
3.1.4. Formulação do Modelo Estatístico	53
3.1.5. Formulação do Modelo Baseado em Redes Neurais	53
3.1.6. Formulação do Modelo Estatístico Utilizando Pca No <i>Dataset</i>	53
3.1.7. Formulação do Modelo Baseado em Redes Neurais Utilizando Pca No <i>Dataset</i>	54
3.1.8. Formulação do Modelo Estatístico Utilizando Rna Para Detecção De <i>Outliers</i> No <i>Dataset</i>	54

3.1.9 Formulação do Modelo Baseado em Redes Neurais Utilizando Pca No Dataset.....	54
3.1.10. Análise de Sensibilidade	54
4. RESULTADOS PARCIAIS	55
4.1. IDENTIFICAÇÃO DAS PROPRIEDADES	55
4.2. PRÉ-PROCESSAMENTO DE DADOS	57
4.3. ANÁLISE ESTATÍSTICA DE VARIÁVEIS	58
4.4. FORMULAÇÃO DO MODELO ESTATÍSTICO	61
4.5. FORMULAÇÃO DO MODELO BASEADO EM REDES NEURAIIS	64
4.6. FORMULAÇÃO DO MODELO ESTATÍSTICO UTILIZANDO PCA NO DATASET	67
4.7. FORMULAÇÃO DO MODELO BASEADO EM REDES NEURAIIS UTILIZANDO PCA NO DATASET	70
4.8. FORMULAÇÃO DO MODELO ESTATÍSTICO UTILIZANDO RNA PARA DETECÇÃO DE OUTLIERS NO DATASET	73
4.9. FORMULAÇÃO DO MODELO BASEADO EM REDES NEURAIIS UTILIZANDO RNA PARA DETECÇÃO DE OUTLIERS NO DATASET	76
4.10. ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS.....	80
5. CONCLUSÕES.....	83
REFERÊNCIAS	85

1. INTRODUÇÃO

No Brasil, as principais fontes de matéria prima na produção de papel e celulose são provenientes das espécies arbóreas de Eucaliptos e Pinus e, através do processo de polpação, é possível obter-se uma polpa de celulose para produção do papel. Este processo de polpação tem grande influência sobre as características e qualidade do produto final. Tendo em vista que o papel é o principal produto da celulose, as empresas possuem a demanda de entregar um produto dentro das normas e padrões de qualidade previstos, o que exige um controle maior no processo de produção, para que se atinja os níveis de qualidade desejados (GOMIDE *et al.*, 2005).

Para satisfazer esta demanda contínua por qualidade exigida pelo mercado, novas tecnologias que possibilitam um maior controle sobre o processo de produção vem sendo alvo de estudos. Uma das principais ferramentas de controle de processos na indústria são os sensores virtuais, que são ferramentas que a partir da leitura de dados fornecidos pelo sistema de produção, possibilita uma análise lógica a partir de um modelo matemático ajustado para o sistema específico, permitindo uma predição sobre a qualidade do produto final (DE CARVALHO *et al.*, 2015).

Os sensores virtuais têm como objetivo a predição de variáveis de interesse do processo em que a leitura das mesmas seja inviável para a indústria, seja pelo alto custo de sensores físicos capazes de medi-la, valor elevado de manutenção ou até mesmo dificuldade técnica para realizá-la, ou a necessidade de métodos não convencionais para a medição desta variável (XUEFENG, 2010). Devido a necessidade das indústrias por um maior controle sobre a produção, este trabalho tem como finalidade o desenvolvimento de um sensor virtual.

O desenvolvimento dos sensores virtuais se dá através de modelos matemáticos preditivos, sendo eles classificados como caixa branca e caixa preta. Os modelos caixa branca são modelos que utilizam de equações fenomenológicas para predição da variável de interesse, também se incluem na definição de caixa branca os modelos lineares. Os modelos caixa preta

possuem uma maior complexidade, pois são os modelos não-lineares, que normalmente conseguem representar o sistema com uma maior fidelidade (ALMEIDA, 2017).

As técnicas de regressão linear multivariável e regressão não-linear por mínimos quadrados parciais multivariável mostram-se útil em qualquer campo da ciência, principalmente em ambiente industrial para o desenvolvimento de sensores virtuais, devido ao fato do tempo e recursos necessários para se desenvolver um modelo com capacidade preditiva de alto grau de excelência a nível industrial, na maioria dos casos não é justificável ou mesmo factível (FACCHIN, 2005).

Outra técnica utilizada para o desenvolvimento dos sensores virtuais são Rede Neural Artificial (RNA) que, simulam um modelo de neurônio biológico, que a partir de sistemas paralelos distribuídos, efetuam cálculos matemáticos nos quais são associados pesos sinápticos que armazenam valores de conhecimento, tornando o modelo confiável, permitindo que seja utilizado tanto em sistemas lineares quanto em não lineares (DUFOUR *et al.*, 2005).

1.1 JUSTIFICATIVAS

Sensores são de fundamental importância para a instrumentação, pois permitem que sistemas de automação e controle tenham a percepção do que ocorre no sistema, fornecendo informações necessárias para o controle da planta e, monitoramento de desempenho do processo. Por ter um papel fundamental na indústria que os sensores vem sendo foco de constante evolução tecnológica, com a finalidade de se reduzir os problemas e potencializar sua atuação.

Os principais problemas dos sensores são erros de medição, disponibilidade e confiabilidade, atrasos de medição, distância do ponto de medição, ambiente de medição, interferência no processo e preço (LOTUFO *et al.*, 2008). Na tentativa de solucionar esses problemas, o estudo de novas tecnologias se faz necessário.

1.2 OBJETIVO GERAL

Considerando a necessidade das empresas por desenvolvimento de novas tecnologias para controle dos processos industriais. Determina-se como objetivo geral deste projeto, o desenvolvimento de um sensor virtual para controle da resistência à tração do produto em uma planta de CTMP (*Chemi-Thermomechanical Pulp*) da empresa de fabricação de papel e celulose KLABIN.

1.3 OBJETIVOS ESPECÍFICOS

Para a construção do sensor virtual, estabeleceu-se objetivos específicos a serem atingidos, sendo eles:

- Investigar as variáveis operacionais mais importantes na definição da resistência à tração do cartão produzido pela CTMP.
- Desenvolver um modelo matemático baseado em redes neurais capaz de prever satisfatoriamente a variável de interesse e que possa ser utilizado como um sensor virtual.

1.4 ORGANIZAÇÃO DA DISSERTAÇÃO

Esta dissertação foi organizada em seis capítulos, sendo este capítulo, onde o tema do trabalho e os objetivos foram introduzidos. O Capítulo 2 é composto pela revisão bibliográfica, onde são revisados os temas pesquisados, tais como, propriedades do papel, processo de produção, métodos de predição, análise estatística, redes neurais artificiais e por fim o desenvolvimento e aplicação de sensores virtuais.

O capítulo 3, apresenta a metodologia utilizada para a realização da análise das variáveis que influenciam a resistência à tração e a construção do

modelo matemático. O capítulo 4 apresenta os resultados obtidos aplicando o modelo desenvolvido e a discussão dos mesmos. Os capítulos 5 e 6 são compostos pelas considerações finais do trabalho e sugestões para trabalhos futuros, respectivamente.

2. REVISÃO BIBLIOGRÁFICA

2.1 PROPRIEDADES DO PAPEL

Utiliza-se no Brasil atualmente diferentes espécies arbóreas de *Eucaliptos* e *Pinus* para a obtenção de celulose e papel a partir da madeira. O processo para obtenção de polpa de madeira, que ocorre antes da produção do papel, pode ser definido como polpação. Este processo utiliza a separação das fibras presentes na madeira através da utilização da energia mecânica, térmica ou química ou até mesmo a combinação das três energias. Durante o processo de polpação química, ocorre a dissolução da lignina, com isto preserva-se a maior parte da celulose e hemicelulose (BASSA *et al.*, 2007).

Os papeis que possuem maior resistência e coloração escura são os papeis oriundos da polpa *kraft*. A polpação que ocorre durante o processo *kraft* é afetada diretamente de pelas características da matéria prima, sendo algumas delas, espécie, qualidade, umidade e densidade, está relacionado também com o licor de cozimento, controle da polpação entre outros parâmetros. A partir destes parâmetros, as empresas de produção de papel e celulose buscam um processo mais uniforme, em que ocorre o mínimo de variação no processo e na qualidade da polpa, obtendo-se um produto que possa ser o mais próximo das especificações finais (delimitadas pelos testes de resistência do papel) (SAMISTRARO *et al.*, 2009).

Antes de passar pela máquina de produção de papel, a polpa obtida passa por um processo de refinação, em que as fibras da polpa sofrem uma ação mecânica, fazendo com que ocorra penetração de água na parede celular, a ruptura de certas ligações de hidrogênio entre a fibra e a fibrilação da parede primária, assim, ocorre uma melhoria no entrelaçamento das fibras além do aumento de flexibilidade. Este processo de refinação influencia diretamente na resistência ao rasgo pois, com o aumento do número de cortes nas fibras, ocorre um decréscimo da resistência ao rasgo, além de um aumento da resistência ao arrebatamento e a tração devido ao aumento da área de contato das fibras. Este processo também influencia nas propriedades de alvura, opacidade e porosidade do papel. Dentre essas variáveis, as que

possuem maior influência sobre as propriedades do produto final estão, o comprimento das fibras, a espessura da parede celular, os danos físicos durante o processo de polpação e deslignificação. Estas propriedades são testadas em ambiente padronizado, e os ensaios podem ser de natureza mecânica, química, elétrica, ótica, propriedades superficiais e permeabilidade. Alguns dos ensaios de natureza mecânica são os que envolvem testes para medir resistência à tração, ao rasgo, a rigidez, e a gramatura (SAMISTRARO *et al.*, 2009).

As características estruturais do papel são espessura, gramatura, a uniformidade e o lado de formação, orientação das fibras, densidade aparente, porosidade e textura, sendo possível observar ou medir diretamente sobre a folha algumas dessas características, já o restante requer a aplicação de testes. Esses testes podem ser divididos em quatro grupos: submissão de esforços, qualidade da superfície, propriedades óticas e permeabilidade a fluidos (SMOOK, 1992).

De acordo com D'Almeida (1998), a resistência à tração está diretamente relacionada à durabilidade do papel para os papéis utilizados em embalagens e outros sujeitos à força de tração, já para os papéis para impressão, isto indica uma probabilidade de ruptura durante a impressão. Esta propriedade do papel está relacionada ao grau de ligação entre fibras, apresentando normalmente maiores índices em polpas com uma quantidade maior de fibra por grama, o que proporciona um contato maior entre as fibras, aumentando o número de ligações.

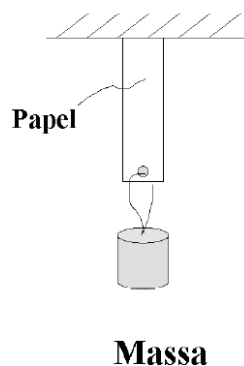
Para a indústria de celulose e papel, as resistências mecânicas são consideradas de grande interesse pois é importante o conhecimento da resistência do papel quando ele estiver em situações em que necessita a aplicação de um esforço sobre ele. O que influencia principalmente a resistência à ação de forças é sua composição fibrosa e sua formação (CIT, 2007).

Segundo Zanuncio (2011), o índice de rasgo pode ser considerado como o trabalho exigido para que o papel comece a rasgar, a uma determinada distância após o rasgo ser aplicado por uma faca adaptada ao

aparelho. Já a resistência à tração pode ser determinada através da relação da força axial necessária para romper uma tira de papel, e a sua largura.

O papel deve ao menos resistir às forças aplicada sobre ele durante o processo de produção e utilização. Um teste que pode ser feito para determinar a resistência à tração do papel, é submetendo um corpo de prova em que a sua largura e comprimento são conhecidos a um esforço de tração crescente de forma uniforme até que ocorra a ruptura do papel (KLOCK *et al.*, 2013). Este exemplo pode ser observado na Figura 1.

Figura 1 – Teste de resistência à tração aplicado sobre o papel.



FONTE: Klock *et al.* (2013).

É possível a utilização o aparelho dinamômetro como um meio de determinar a resistência à tração (Figura 2). Também conhecido como comprimento de auto-ruptura (CAR), a resistência à tração através da Equação 2.1.

$$CAR = \frac{C_r \times 1000000}{G \times L} \left(\frac{kgf \times m}{g} \right) \quad (\text{Equação 2.1})$$

Em que C_R é a carga de ruptura em kgf medida pelo dinamômetro, G a gramatura da amostra em g/m² e L a largura do corpo de prova em mm.

Figura 2 – Dinamômetro utilizado para calcular a resistência à tração.



FONTE: Klock *et al.* (2013).

2.2 PROCESSO DE PRODUÇÃO DE UMA PLANTA CTMP

A planta de produção de polpa químio termo mecânico da madeira é responsável por abastecer a produção da Máquina de Papel 9, em Telêmaco Borba, e esta polpa tem como objetivo dar volume e espessura ao cartão, compondo a camada do meio, sendo utilizado apenas madeira da espécie Eucalipto para a sua produção.

A madeira é preparada para que se atenda às especificações do processo de produção, algumas dessas especificações são espécie, limpeza e dimensões. Posteriormente, o cavaco da madeira que entra na planta CTMP passa por uma lavagem para remoção de areia e materiais similares. Aplica-se vapor saturado no cavaco, passando por uma prensa, seguido por um ligeiro pré-tratamento químico (impregnação com peróxido de hidrogênio, ácido dietilenotriaminopentacético e hidróxido de sódio) dentro de um silo, tendo como objetivo amolecer a estrutura lenhosa da madeira.

Após o pré-tratamento o cavaco passa por um processo mecânico intenso para a separação das fibras, sob pressão, passando por uma série de refinadores de disco, fazendo com que se obtenha um número elevado de finos, o que faz da pasta produzida uma pasta similar a pasta mecânica convencional.

A pasta passa por uma etapa de separação, e as fibras que não atingiram um tamanho desejado são enviadas novamente para refinação para um melhor aproveitamento da matéria prima. A pasta que atingir os padrões especificados é lavada para a remoção dos químicos presentes. Após a lavagem a água drenada e passa por um processo de tratamento para que possa ser reaproveitada. A pasta segue da lavagem para um processo de prensagem para a redução de umidade da pasta, ficando pronta para suprir a demanda da Máquina de Papel 9. A Figura 3 representa o fluxograma da planta de produção CTMP.

2.3 MODELOS DE PREDIÇÃO

Uma ferramenta muito utilizada para a determinação de resultados sem a necessidade de realização experimental para conferência são os modelos matemáticos. Eles são a representação abstrata da realidade por equações. A obtenção de um modelo que represente fielmente o sistema é importante pois possibilita a simulação do sistema para obtenção de resultados antes mesmo de se investir recurso para a construção do sistema. Os modelos possuem a finalidade de reproduzir as respostas do sistema real, independente do evento que possa ocorrer ao longo do tempo (GARCIA, 2005).

Modelos em caixa branca são aqueles que se utilizam de equações físicas ou químicas para obtenção de seus resultados, descrevendo o comportamento do sistema. Muitas vezes, a complexidade do sistema e o número de equações necessárias para sua representação dificulta a obtenção do modelo, o que as vezes exige medição de dados de saída do sistema, a partir da variação da entrada (GARCIA, 2005).

São considerados modelos lineares, aqueles que as variáveis que os definem estiverem representadas em equações sob a forma de funções lineares. Entretanto os modelos fenomenológicos podem ser descritos através das leis fundamentais da física e da química, onde equações descrevem o sistema (CASILLO, 2009).

Se o sistema for representado por equações não lineares, os modelos em caixa branca se tornam insuficientes para descrever o sistema, sendo melhor representado em modelo conhecido como caixa preta, por abrangerem sistemas mais complexos. Os modelos não lineares representam o sistema com maior fidelidade, mas possuem uma complexa identificação, fazendo com que eles necessitem de uma maior atenção do que os sistemas lineares (ALMEIDA, 2017).

Na indústria química, os modelos fenomenológicos não-lineares enfrentam um grande problema, que é seu alto custo para desenvolvimento e validação e, a etapa de modelagem é responsável pela maior parte do custo

em projetos de desenvolvimento de sistemas de controle avançado (HUSSAIN, 1999).

Para resolver este problema, uma alternativa para o desenvolvimento desses sistemas, podem-se utilizar inferidores capazes de assimilar o comportamento do processo através da análise de seus dados históricos, desenvolvendo uma estrutura empírica em que seus parâmetros devem ser estimados. Estes estimadores foram favorecidos devido ao enorme avanço dos equipamentos na área de tecnologia digital, por serem baseados na manipulação de grandes quantidades de dados e cumprirem rotinas trabalhosas (ZYNGIER, 2000).

Existem vários métodos de inferência com capacidade de determinar valores de variáveis e parâmetros de interesse em tempo real. A definição da metodologia a ser utilizada depende de fatores que variam caso por caso, em que modelos caixa preta podem ser desenvolvidos em um espaço mais curto de tempo, já os modelos caixa branca, apesar de necessitarem de mais tempo, possuem uma maior capacidade de extrapolação (ZYNGIER, 2000). A Tabela 1 representa uma síntese dos métodos aplicados em análise multivariável.

Tabela 1 – Visão simplificada do campo de análise multivariável.

	Comunidade	Palavras Chave	Época
MLR <i>Multiple Linear Regression</i>	Todas	Ajuste	1800
PCA <i>Principal Component Analysis</i>	Quase todas	Compressão de dados	1930
PLS <i>Partial Least Squares</i>	Quimiometria	Calibração de modelos	1970
CCA <i>Canonical Correlation Analysis</i>	Estatísticos	Variáveis canônicas	1960
FA <i>Factor Analysis</i>	Ciências Sociais	Fatores ocultos	1930
ICA <i>Independent Canonical Analysis</i>	Redes Neurais	Reconhecimento de padrões	1990
SSI <i>SubSpace Identification</i>	Teóricos de Controle	Identificação de sistemas	1990
NNR <i>Neural Network Regression</i>	Quase todas	Aprendizagem	1980

FONTE: FACCHIN, 2005

2.4 MODELOS DE PREDIÇÃO ESTATÍSTICOS

2.4.1 Análise de Variáveis

A parte de seleção de variáveis na construção do modelo matemático é de extrema importância, independentemente de sua tipologia, juntamente com o pré-processamento dos dados, sendo os principais responsáveis pelo sucesso do modelo desenvolvido (FACCHIN, 2005).

A dependência entre variáveis faz com que uma variável carregue consigo informações sobre a outra. Esta dependência pode ser tanto forte quanto fraca e, de sentidos iguais ou opostos. A correlação de Pearson descreve a medida da dependência linear entre as variáveis, já a correlação de Spearman é utilizada para determinar a correlação não-linear entre as variáveis (CASELLA *et al.*, 2002).

ARANHA (2016) descreve a correlação de Pearson na Equação 2.2.

$$r_p(X, Y) = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum(x_i - \bar{x})^2)(\sum(y_i - \bar{y})^2)}} \quad (\text{Equação 2.2})$$

Onde o sinal de coeficiente indica o sentido da relação e a magnitude do coeficiente indica a intensidade da relação linear entre as variáveis. Quanto mais próximo o valor do módulo de r_p for de 0, mais fraca é a correlação e, quanto mais próximo de 1 mais forte é a correlação linear das variáveis.

A correlação de Spearman é a aplicação do coeficiente de Pearson em um conjunto de dados após uma classificação dos mesmos. Quando os dados estão em *ranks* os *outliers* do conjunto não provocam uma grande alteração no comportamento do coeficiente de correlação. Os *ranks* da correlação de Spearman classificam tanto as variáveis positivas quanto as negativas, independente do tamanho do grupo de variáveis. As diferenças entre os ranks positivos e negativos é gravada. Em um conjunto de dados de tamanho n , os valores dos dados A_i e B_i são convertidos para os ranks rgA_i e rgB_i , assim possibilitando o cálculo da correlação de Spearman através da correlação de

Pearson como descrito na Equação 2.3 (KUMAR *et al.*, 2018; PUTH *et al.*, 2015).

$$r_s = \rho_{rgA,rgB} = \frac{cov(rgA,rgB)}{\sigma_{rgA} \cdot \sigma_{rgB}} \quad (\text{Equação 2.3})$$

Em que ρ é a correlação de Pearson entre os ranks das variáveis, $cov(rgA,rgB)$ a covariância do rank das variáveis e, σ_{rgA} e σ_{rgB} são o desvio padrão dos ranks das variáveis.

Se todos os ranks das variáveis são distintos, utiliza-se a Equação 2.4.

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} \quad (\text{Equação 2.4})$$

Onde $d_i = rg(A_i) - rg(B_i)$ é a diferença significativa entre os dois ranks de variáveis e, n é o número total de observações.

2.4.2 Regressão Linear Multivariável

Também conhecida como MLR (*Multiple Linear Regression*), a regressão linear multivariável é uma complementação da regressão linear simples em casos que necessitam de uma abordagem diferente, pelo fato de possuírem mais de uma variável de entrada. O método tem como objetivo analisar a relação linear entre um conjunto de variáveis X , que é composto por m vetores de comprimento n e, a variável resposta Y , que é um vetor de comprimento n (GELADI *et al.*, 1986). Essa relação pode ser expressa através da Equação 2.5.

$$y = b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \dots + b_m \cdot x_m + e \quad (\text{Equação 2.5})$$

Ou de forma simplificada:

$$y = \sum_{j=1}^m b_j \cdot x_j + e \quad (\text{Equação 2.6})$$

As Equações 2.5 e 2.6 só se aplicam a casos em que n é igual a 1, já para os casos que apresentarem um n maior que 1, a equação terá uma característica de matriz (GELADI *et al.*, 1986). Observa-se essa notação matricial na Equação 2.7.

$$y = X.b + e \quad (\text{Equação 2.7})$$

Onde y é o vetor com os valores de saída (resposta) com n linhas e 1 coluna, X é a matriz composta por variáveis de entrada possuindo n linhas e m colunas, b é um vetor composto por parâmetro responsável pela obtenção de uma estimativa de saída a partir da entrada, e e que é um vetor contendo os resíduos, representada pela Equação 2.8.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1m} \\ X_{21} & X_{22} & X_{23} & \dots & X_{2m} \\ X_{31} & X_{32} & X_{33} & \dots & X_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & X_{n3} & \dots & X_{nm} \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix} \quad (\text{Equação 2.8})$$

2.4.3 Regressão por Mínimos Quadrados Parciais

O método dos mínimos quadrados parciais, em inglês PLS (*Partial Least Squares*), foi introduzido por Wold em 1975, e é utilizado para determinação dos pontos mínimos de uma função que significa o desvio estimado na busca pelo ajuste. É geralmente utilizada na genética e na quimiometria, onde o número de observações é inferior ao número covariáveis, e possuem alta correlação entre si (ALMEIDA, 2015).

É considerado útil para o desenvolvimento de equações capazes de prever o comportamento de um sistema com grande número de variáveis de entrada e um conjunto de dados amostral pequeno, formando componentes que absorvam grande parte da informação disponível nas variáveis de entrada (X_1, X_2, \dots, X_n), para conseguir prever a resposta das variáveis de saída (Y_1, Y_2, \dots, Y_j). Possui similaridade com a Regressão por Componentes Principais (PCR), diferindo no fato de considerar as variáveis respostas, enquanto o PCR não as leva em conta (SILVEIRA, 2014).

As principais vantagens do PLS envolvem a capacidade de regressão com múltiplas variáveis de saída, não é afetado pela multicolinearidade das variáveis de entrada, produzindo fatores com alto poder de predição (MORELLATO, 2010).

WANG (2019), descreveu a regressão PLS em seu trabalho. Considerando que $Y \in R^{n \times q}$, $X \in R^{n \times p}$ correspondem a matriz resposta e a matriz de predição respectivamente. O núcleo deste método está na hipótese da decomposição das Equações 2.9 e 2.10:

$$X = T.P^T + E \quad (\text{Equação 2.9})$$

$$Y = U.Q^T + F \quad (\text{Equação 2.10})$$

Onde $T_{n \times h} = (t_1, t_2, \dots, t_h)$ e $U_{n \times h} = (u_1, u_2, \dots, u_h)$ são variáveis latentes de dimensão h , $P_{p \times h}$ e $Q_{q \times h}$ são matrizes de coeficientes, $E_{n \times h}$ e $F_{n \times h}$ são matrizes residuais. O objetivo do método é encontrar os pesos w e c para X e Y respectivamente, maximizando a covariância entre a pontuação latente de t e u .

O algoritmo NIPALS (Nonlinear Iterative Partial Least Squares) é o método de estimação mais comum para a solução de PLS, que procura as projeções de direções otimizadas resolvendo o seguinte problema:

$$\max [cov(Xw, Yc)]^2, s. t. \|w\| = 1, \|c\| = 1;$$

WOLD (1975) descreveu os passos do algoritmo NIPALS:

1. Inicialização do vetor de pontuação do vetor u
2. Cálculo de w e padronização do mesmo. $w = \frac{X^T u}{u^T u}, w = \frac{w}{\|w\|}$
3. Cálculo de t
4. Cálculo de c e padronização do mesmo. $c = \frac{Y^T t}{t^T t}, c = \frac{c}{\|c\|}$
5. Cálculo de u
6. Repetir do passo 2 ao 5 até convergir
7. Deflação de X e Y . $X = X - t.p^T, Y = \frac{Y - t.t^T Y}{t.t^T} = Y - tc^T$

2.4.4 Análise de *Cluster*

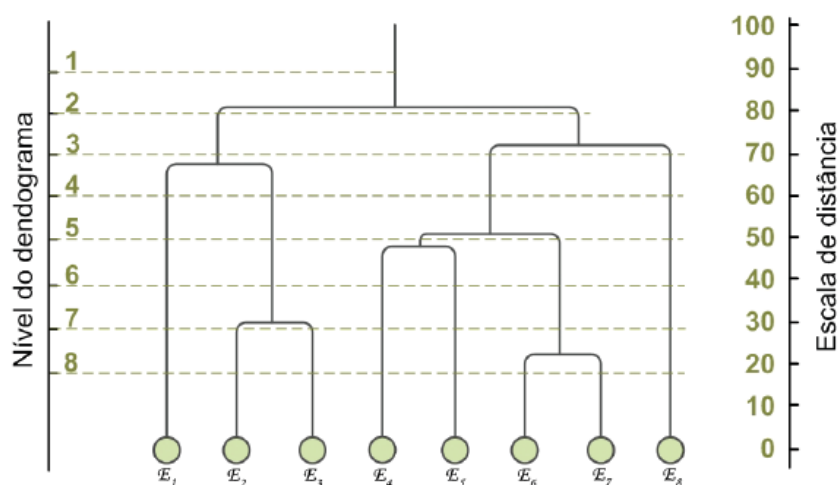
A análise de *cluster*, também conhecida como análise de agrupamento, são diversos métodos matemáticos, que possuem a finalidade de classificar as observações dentro de um espaço p -dimensional, de uma maneira em que eles se dividam em um número de grupos de forma conveniente para que seja possível observar a similaridade presente em cada grupo. Esta técnica é utilizada a fim de descobrir estruturas em conjunto de dados e separá-los em agrupamentos (*clusters*), em que seus elementos apresentem características parecidas (KLEN; LANA, 2014).

JAIN (2010) definiu que o objetivo da análise de *cluster* dividir o conjunto de dados de uma maneira que ocorra uma homogeneidade no grupo, onde os dados que pertencem ao mesmo grupo possuem a maior similaridade, heterogeneidade entre os grupos, em que dados pertencentes a diferentes grupos são o mais diferente possível dos outros grupos, e que os agrupamentos são regiões de alta densidade de dados, e que são separados entre eles por regiões de baixa densidade de dados.

Os métodos de análise de *cluster* necessitam de uma maneira para medir a similaridade entre os elementos, esta medida é conhecida como função distância, entre essas medidas temos a distância euclidiana, que é uma distância geométrica no espaço multidimensional, distância euclidiana quadrática, distância de Manhattan, distância de Chebychev (HALKIDI *et al.*, 2001).

Os algoritmos utilizados para o agrupamento podem ser divididos em algoritmos hierárquicos e algoritmos de particionamento. Os algoritmos hierárquicos formam estruturas encadeadas de uma forma que os elementos com maior similaridade sejam agrupados de uma maneira mais próxima, podendo ser representado graficamente por um dendograma (diagrama de árvore), em que cada ramo representa um elemento e a raiz é representada pelo agrupamento de todos os elementos. A Figura 4 demonstra um exemplo de dendograma.

Figura 4 – *Clustering* representado por meio do Dendograma.



FONTE: Metz (2006).

A partir do dendograma e do conhecimento sobre o conjunto de dados é possível determinar uma distância, para que assim o agrupamento dos dados seja formado. Dentre estes métodos está o método aglomerativo que pode ser subdivido em diversos métodos, sendo eles o método de ligação por vizinhos mais próximos, o método de ligação por vizinhos mais distantes, o método de ligação por média, o método de ligação por centroide, o método de ligação por mediana, o método de ligação de Ward.

Já os algoritmos de particionamento são classificados de acordo com a forma em que eles são capazes de lidar com as incertezas que aparecem devido a sobreposição dos agrupamentos, sendo eles divididos em rígido (*hard*) e nebulosos (*fuzzy*). Os algoritmos rígidos dividem o conjunto de dados de uma maneira que o critério de divisão está relacionado com a distância dos pontos em relação ao centro do grupo, fazendo com que ele seja membro exclusivamente daquele grupo. Já os algoritmos *fuzzy* permite que um elemento seja membro de mais de um grupo, se tiver o necessário a inclusão ao grupo, fazendo com que se tenha uma divisão mais realista dos elementos (KLEN; LANA, 2014).

2.4.5 Análise dos Componentes Principais

Análise dos Componentes Principais (Principal component analysis - PCA), é uma técnica multivariada que pode ser aplicada a um conjunto de dados em que, a partir da combinação linear das variáveis do conjunto analisado, permite realizar uma redução de dimensionalidade de matrizes. Ao aplicar-se esta técnica a uma matriz contendo muitas variáveis de entrada, forma-se um novo conjunto de variáveis que são ortogonais entre si, não correlacionando com os dados originais, e suas primeiras componentes do novo conjunto de dados, que são denominadas componentes principais, são responsáveis por explicar a maior parte da variância total que existe no conjunto de dados, fazendo com que sejam capazes de representá-los (LOPES-HIDALGO *et al.*, 2018; ROTARU, 2014).

É um método que tem como objetivo maximizar a taxa de redução de variância. Devido aos problemas de seleção e extração de características que são comuns em reconhecimento estatístico de padrões. Esta seleção de característica é um processo em que o espaço de dados se transforma em espaço de características, a fim de reduzir o número de características efetivas e ao mesmo tempo conservando o conteúdo das informações dos dados originais, reduzindo a dimensionalidade dos dados enquanto se tem uma transformação representativa e compacta. Este método consiste em transformar um vetor $X \in R^m$ em um novo vetor $X \in R^n$, em que $n \leq m$, onde se projeta X nas n direções ortogonais de maior variância, que são os componentes principais (RIBEIRO, 2007).

A matriz original é decomposta e projetada em duas partes, as amostras são classificadas na parte de pontuação e os descritores em termos de separação da amostra na parte de carregamento de informação, seguindo o eixo dos componentes principais (DJELLOULI *et al.*, 2018). Os detalhes do PCA podem ser resumidos a partir das seguintes equações:

A Equação 2.11 representa a matriz original de múltiplas características.

$$X^0 = \begin{bmatrix} X_{11} & \cdots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{m1} & \cdots & X_{mn} \end{bmatrix} \quad (\text{Equação 2.11})$$

Onde m é o número de ensaios e n é o número de características de qualidade. Realiza-se uma padronização dos dados, obtendo a Equação 2.12.

$$X = \begin{bmatrix} \bar{X}_{11} & \cdots & \bar{X}_{1n} \\ \vdots & \ddots & \vdots \\ \bar{X}_{m1} & \cdots & \bar{X}_{mn} \end{bmatrix} \quad (\text{Equação 2.12})$$

Através da Equação 2.13 é possível obter a Equação 2.14, que representa a matriz de correlação de coeficientes.

$$r_{kj} = \frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j) \times (X_{ik} - \bar{X}_k)}{\sqrt{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \times \sum_{i=1}^n (X_{ik} - \bar{X}_k)^2}} \quad (\text{Equação 2.13})$$

$$R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix} \quad (\text{Equação 2.14})$$

Onde \bar{X}_k e \bar{X}_j são os valores médios de X_k e X_j . A determinação dos autovalores e autovetores estão representados na Equação 2.15.

$$(R - \lambda_j) \times V_j = 0 \quad (\text{Equação 2.15})$$

Em que λ_j representa os autovalores, com $\sum_{j=1}^n \lambda_j = n$, e $j = 1, 2, 3, \dots, m$; e, $V_j = (a_{j1}, a_{j2}, a_{j3}, \dots, a_{jn})$ são os autovalores correspondentes aos autovetores. A variação explicada é representada pelas Equações 2.16 e 2.17.

$$b_j = \frac{\lambda_j}{\sum_{j=1}^n \lambda_j} \quad (\text{Equação 2.16})$$

$$a_p = \frac{\sum_{j=1}^p \lambda_j}{\sum_{j=1}^n \lambda_j} \quad (\text{Equação 2.17})$$

Onde b_j corresponde a taxa de contribuição do j-ésimo autovalor, enquanto a_p representa a taxa de contribuição acumulativa do p-ésimo autovalor. Assim podemos formular a Equação 2.18.

$$Y_{mj} = \sum_{i=1}^n X(i)_m \times V_{ij} \quad (\text{Equação 2.18})$$

Em que Y_{mj} é a j-ésimo dado formulado a partir dos autovetores. Assim, a partir do cálculo da taxa de contribuição e da acumulação da taxa de contribuição para a variância total, os componentes principais são ordenados, onde o primeiro componente possui a maior variância possível e cada componente seguinte são determinados com a propriedade de que eles são ortogonais ao primeiro componente principal e que é a maior variância possível (DJELLOULI *et al.*, 2018).

A seleção de quais componentes principais utilizar em seu novo conjunto de dados é uma das principais discussões existentes nesse assunto. Um dos métodos mais utilizados é o critério de KAISER (1958), em que ele propõe considerar apenas os componentes principais em que os autovalores são superiores a 1, ou seja, são os componentes que explicam a maior parte da variação no conjunto de dados.

Outro método de seleção dos componentes principais comumente utilizado são pela observação da proporção de explicação da variância total, que em muitos casos, seleciona-se componentes suficientes para explicar pelo menos 80% da variação total do modelo, escolhendo o primeiro componente principal, que é o de maior importância, responsável pela maior variância, seguido pelo segundo componente principal de maior variância, sucedendo até o componente seguinte responsável pela maior variância até atingir a porcentagem desejada (JOHNSON; WICHERN, 1998; HONGYU *et al.*, 2015). JOLLIFFE (1973) adotou um critério em que ele seleciona um número componentes principais igual ao número de componentes cujo o autovalor é menor que 0,7 assim, descartando todos os outros componentes principais.

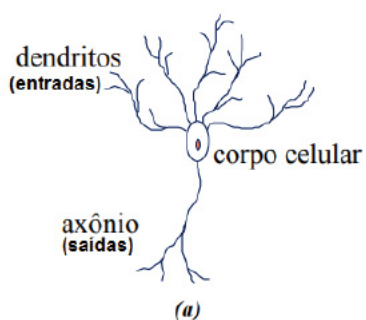
É possível selecionar componentes principais através de uma análise em um diagrama de autovalores, mantendo os componentes em que os autovalores estão situados acima do ponto de mudança ruptura da inclinação da curva da função que relaciona a ordem e os autovalores, analisando a porcentagem de explicação do componente principal, onde se os componentes apresentarem o mesmo grau de variabilidade, não havendo

motivo para conservar um e não conservar o de valor proporcional e, uma grande diminuição do autovalores entre um componente e seu sucessor, significa uma conservação da interpretação dos componentes anteriores (RIBEIRO, 2007).

2.5 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNA) se utilizam da simulação de neurônios biológicos, como mostrado na Figura 5, por meio de funções matemáticas paralelamente distribuídas. Os neurônios biológicos têm como característica, a habilidade de aprender através da experiência, sendo capaz de desenvolver suas próprias regras para seu perfeito funcionamento (HAYKIN, 2001). Elas são uma forma de computação não-algorítmica com forma estrutural semelhante à de neurônios biológicos. Efetuam cálculos matemáticos nos quais são associados pesos sinápticos que armazenam valores de conhecimento, assim, os sistemas paralelos distribuídos formam um modelo confiável para sistemas lineares e não lineares (BRAGA *et al.*, 2000).

Figura 5 – Representação de um neurônio biológico.

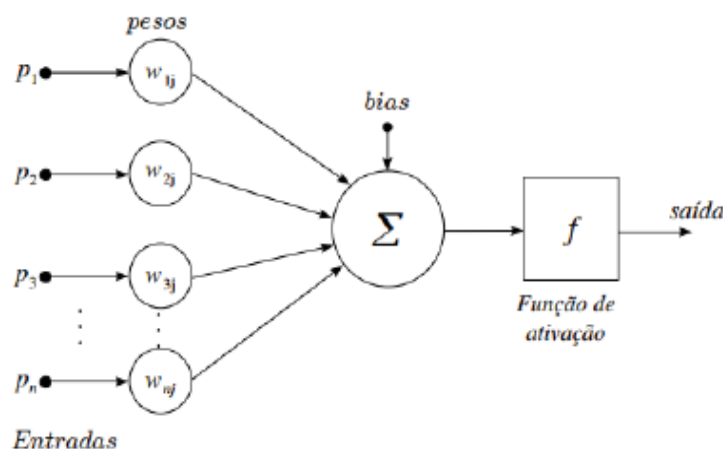


FONTE: Poppi et al. (2001).

Os neurônios artificiais podem ser representados graficamente de uma maneira similar ao neurônio biológico, como pode ser observado na Figura 6,

este possui como dendritos as entradas do sistema, que são multiplicadas aos respectivos pesos sinápticos, seguido por uma unidade de processamento de informações, para a saída (axônio) de uma função de ativação que restringe a amplitude da saída do neurônio artificial (HAYKIN, 2001).

Figura 6 – Representação de um neurônio artificial.



FONTE: Haykin (2001).

No modelo não linear representado acima, é possível observar três elementos básicos:

- **Sinapses:** É caracterizada por pesos ou forças próprias. O Sinal de entrada p_n que está conectado ao neurônio k é multiplicado pelo peso sináptico w_{kj} .
- **Junção aditiva:** Responsável pela soma ponderada dos sinais de entrada
- **Função de Ativação:** Responsável pela restrição da amplitude da saída de um neurônio, limitando o intervalo de amplitude do sinal, podendo ser linear ou não-linear.

O modelo acima também apresenta um *bias*, que possui a função de aumentar ou diminuir a entrada líquida da função de ativação (STEIN, 2010).

Para se estruturar uma rede neural deve-se levar em conta alguns fatores, como a quantidade de entradas e saídas, a dependência da rede com as saídas e entradas anteriores, o grau de não linearidade do sistema, o número de neurônios e de camadas de neurônios a serem utilizadas, a interação entre os neurônios e a melhor função de ativação para o sistema. Um sistema que esteja bem estruturado tem a capacidade de representar fielmente um sistema não linear, contanto que seja treinado e validado de maneira correta (BRAGA *et al.*, 2000).

Se representarmos os sinais de entrada por $\{x_1, x_2, x_3, \dots, x_n\}$ advindos do meio externo com valores assumidos pelas variáveis de uma aplicação específica e, os pesos sinápticos representados por $\{w_1, w_2, w_3, \dots, w_n\}$ que irão ponderar as variáveis de entrada, é possível representar matematicamente um neurônio k através do seguinte par de equações (2.19) e (2.20) propostas por McCulloch e Pitts (1969):

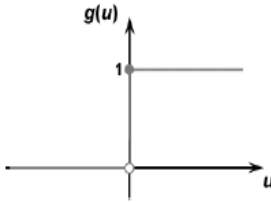
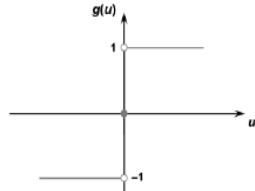
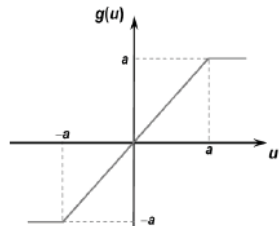
$$u = \sum_{i=1}^n w_i \times x_i - \theta \quad (\text{Equação 2.19})$$

$$y = g(u) \quad (\text{Equação 2.20})$$

Sendo o *bias* representado por θ , o potencial de ativação u que é o resultado da soma dos valor produzido pelo combinador linear e a limiar de ativação, função de ativação g que limita a saída do neurônio k dentro de um intervalo de valores de interesse do sistema, e o sinal de saída y com o valor final produzido pelo neurônio derivado dos valores de entrada especificados (BATISTA, 2012).

Os tipos de função mais utilizadas são as funções limiar, linear, linear por partes e sigmoidal. As saídas, que são representadas pelas equações presentes nas Tabelas 2 e 3, podem ser divididas em funções de ativação parcialmente diferenciáveis e funções de ativação totalmente diferenciáveis.

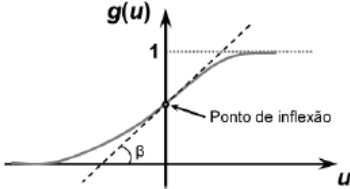
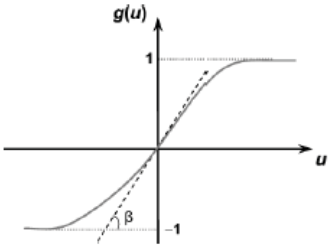
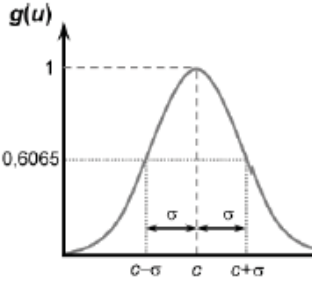
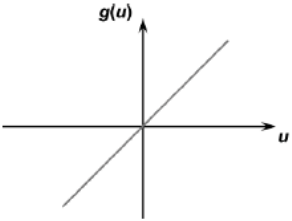
Tabela 2– Principais funções de ativação parcialmente diferenciáveis.

Função	Equação	Representação Gráfica
Função degrau	$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases}$	
Função degrau bipolar ou função sinal	$g(u) = \begin{cases} 1, & \text{se } u > 0 \\ 0, & \text{se } u = 0 \\ -1, & \text{se } u < 0 \end{cases}$	
Função rampa simétrica	$g(u) = \begin{cases} a, & \text{se } u > a \\ u, & \text{se } -a < u < a \\ -a, & \text{se } u < -a \end{cases}$	

FONTE: Batista, 2012.

Como observa-se na Tabela 2, nas funções de ativação parcialmente diferenciáveis, as derivadas de primeira ordem das funções são inexistentes. Já as funções de ativação totalmente diferenciáveis, são funções em que suas derivadas de primeira ordem existem e se tem conhecimento de todos os pontos de seu domínio, como observado na Tabela 3.

Tabela 3 – Principais funções de ativação totalmente diferenciáveis.

Função	Equação	Representação Gráfica
Função Sigmoidal (logística)	$g(u) = \frac{1}{1 + e^{-\beta u}}$	
Função tangente hiperbólica	$g(u) = \frac{1 - e^{-\beta u}}{1 + e^{-\beta u}}$	
Função gaussiana	$g(u) = e^{\frac{-(u-c)^2}{2\sigma^2}}$	
Função linear	$g(u) = u$	

FONTE: Batista, 2012.

A função de ativação sigmoide é a mais utilizada em RNAs do tipo MLP (Multilayer Perceptron) devido ao fato de ser a que mais se aproxima de um neurônio biológico (STEIN, 2010).

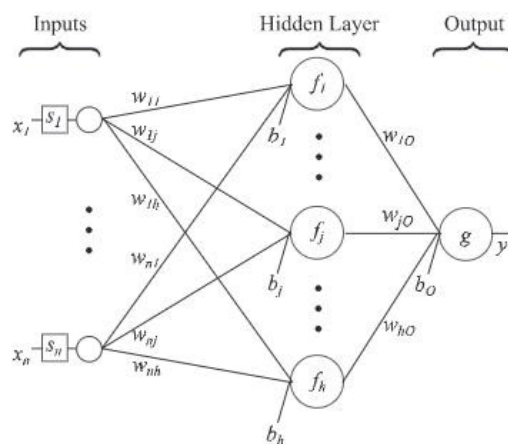
2.5.1 Tipos de Arquiteturas, Treinamento de Redes Neurais e Validação

O desenvolvimento de uma RNA passa pelas seguintes etapas: seleção da arquitetura em que ocorre a definição da quantidade de nós de entrada, neurônios escondidos e número de saída; o treinamento da RNA através de um conjunto de dados do processo; e validação da rede treinada, podendo ser feita através da validação cruzada, em que se avalia as respostas da rede em relação a dados não utilizados durante o treinamento (HAYKIN, 2001).

A arquitetura representa a maneira em que a RNA está estruturada e é de grande importância, pois é um limitante para o tipo de problema a ser tratado pela RNA. A estrutura das redes neurais é feita de forma heurística, sendo composta por três partes, sendo elas a camada de entrada, as camadas ocultas e a camada de saída. Os tipos de arquiteturas são as de rede alimentada adiante (*Feedforward*) de uma única camada, redes alimentadas adiante de múltiplas camadas (*Multilayer Feedforward Network*) e redes recorrentes (HAYKIN, 2001).

Os modelos do tipo Feedforward estão estruturados de modo que todas as saídas dos neurônios de uma camada estejam conectadas com todos os neurônios da camada posterior assim, obedecendo a direção entrada → saída e, não tendo interação entre os neurônios da mesma camada (FLECK, 2016). Este tipo de estrutura pode ser observado na Figura 7.

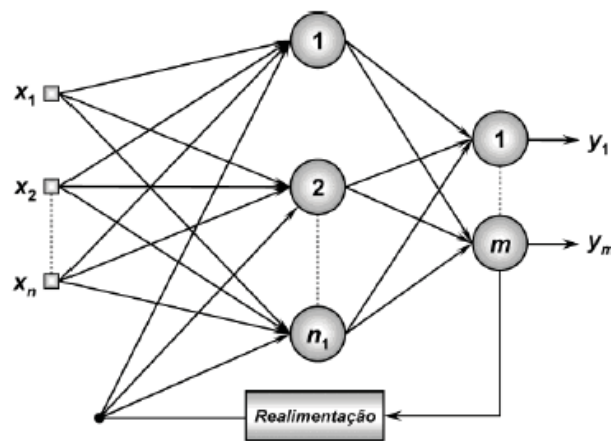
Figura 7 – Arquitetura *Feedforward* com uma camada oculta.



FONTE: Matias et al. (2014).

Já os modelos do tipo Recorrente, ao contrário da estrutura dos modelos *Feedforward*, não possui um único sentido para o fluxo de sinais entre neurônios ou entre camadas, podendo ter conexões entre os elementos de processamento de mesma camada, das camadas anteriores e posteriores, configurando uma rede com realimentação, em que um neurônio pode ser diretamente ou indiretamente retroalimentado pela saída (FLECK, 2016). Este tipo de estrutura pode ser observado na Figura 8.

Figura 8 – Neurônio artificial com uma camada com estrutura recorrente.



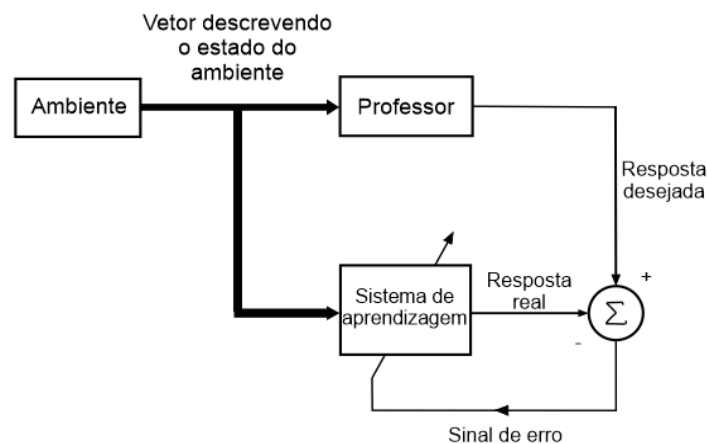
FONTE: Batista (2012).

A estrutura de RNA mais utilizada é a Perceptron de múltiplas camadas, em que possuem múltiplas camadas de unidades básicas de processamento do tipo Perceptron, possuem treinamento supervisionado, em que o algoritmo de treinamento visa ajustar os pesos da rede para obter respostas (saída) desejadas a partir de um conjunto conhecido de vetores de entrada. O algoritmo utilizado para o treinamento é o *backpropagation* (ALMEIDA, 2017).

O Algoritmo consiste em apresentar a rede um conjunto de treinamento, contendo pares de vetores de entrada e saída conhecidos. O vetor de entrada é propagado através das camadas da rede, e o vetor de saída da rede é então comparado ao vetor de saída desejada assim, calculando o erro de saída da rede, que é a diferença entre as saídas. Este erro é retropropagado (*backpropagation*) a rede, como um sinal de erro, para que

ocorra um ajuste nos pesos, a fim de reduzir o erro para a próxima iteração. O processo de retropropagação se repete para todos os pares de treinamento (entrada/saída) da rede, até que se atinja a condição de parada do treinamento, com um erro aceitável (AGUIAR, 2010). Pode-se observar os passos de treinamento através da Figura 9.

Figura 9 – Estrutura do treinamento supervisionado *backpropagation* de uma rede de múltiplas camadas.



FONTE: Haykin (2001).

Pode-se validar a RNA através do erro quadrático médio (RMSE-*Root Mean Square Error*), que é representado pela Equação 2.21 (ALMEIDA, 2017).

$$RMSE = \frac{\sqrt{\sum_{k=1}^n (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^n ((y(k) - \bar{y}(k))^2)}} \quad (\text{Equação 2.21})$$

Em que $\hat{y}(k)$ é a resposta da RNA treinada, $y(k)$ é a saída real e, $\bar{y}(k)$ o valor médio do sinal de saída. Se RMSE obtiver valores maiores que 1, significa que o modelo gerado pela RNA não atende bem o sistema, e ele será melhor representado quanto mais próximo o RMSE for de 0 (AGUIRRE, 2004).

2.5.2 Modelos Baseados em Redes Neurais Aplicados na Indústria de Celulose e Papel

As redes neurais ganharam espaço nos processos industriais devido as suas seguintes propriedades: não-linearidade; mapeamento de entrada; adaptabilidade; e generalização. Resolvendo problemas mais complexos presentes na indústria e, devido a sua arquitetura, possuem maior desempenho comparado a outros modelos.

VIEIRA (2003) através de RNA, desenvolveu modelos empíricos de uma máquina de produção de cartão de embalagem, capazes de prever propriedades físicas do cartão, como resistência à flexão, resistência à delaminação, resistência à absorção de ácido láctico e peróxido de hidrogênio. Desenvolveu-se modelos não-lineares quase-estáticos híbridos, se aproveitando de conhecimento disponível, colaborando para modelar os aspectos não conhecidos do processo, e a RNA se mostrou muito robusta, conseguindo capturar as não-linearidades do processo.

RIBEIRO (2007) desenvolveu um modelo de RNA capaz de prever o grau de cozimento ($\#Kappa$) que ocorre em um digestor contínuo de uma planta de celulose. Com o auxílio da tecnologia FT-NIR (*Fourier Transform Near Infrared*) foi possível identificar os cavacos de madeira que entraram no processo de cozimento, e utilizou-se de uma técnica de análise multivariada PCA (*Principal Components Analysis*) com o objetivo de reduzir a dimensionalidade dos padrões de entrada do modelo. Entradas que seriam as variáveis do processo e as análises feitas pelo FT-NIR nos cavacos da madeira. Através de testes utilizando métodos de otimização paramétrica não-linear no treinamento da RNA, resultando em uma boa adequação na utilização dessas técnicas, e o modelo foi capaz de prever a variável $\#Kappa$, que por meio de ações operacionais mais proativas, ocorreu a redução da variabilidade da mesma. A RNA ofereceu coeficientes de correlação satisfatórios entre fatores preditos e efetivamente medidos no processo.

STEIN (2010) aplicou MAG (Modelos Aditivos Generalizados) e RNA em dados que foram observados em escala industrial durante anos, para

analisar a influência de algumas variáveis medidas, tanto da madeira quanto do processo de produção, sobre a produção do digestor e carga alcalina. Os resultados mostraram que o MAG é uma boa opção para representar os fenômenos que ocorrem na indústria de celulose, onde as variáveis apresentam alta variabilidade e que não ocorre um rigoroso controle, diferente dos dados provenientes de delineamentos experimentais. O uso da RNA para estimar a produção do digestor e para carga alcalina obteve resultado satisfatório, em que a correlação entre os dados reais e estimados ficaram acima de 88% e 60% respectivamente.

GARCIA (2010) a fim de avaliar a qualidade de impressão de cartões com três camadas produzidos pela MP-9 da Klabin, utilizou-se de modelos matemáticos baseados em RNA para determinar os parâmetros de processo que mais influenciam a propriedade de marmorização (*mottling*), a fim de reduzir a variabilidade desta propriedade para mantê-la dentro dos limites especificados. Os resultados obtidos indicaram que a marmorização sofre influência da temperatura das capotas secadoras de revestimento.

ALMEIDA (2008) utilizou-se de uma metodologia que se baseia na técnica HHM (*Hidden Markov Model*), para que fosse possível fazer a detecção de situações anormais em caldeiras de recuperação química, no seu estado inicial, para evitar uma condição indesejável. Foi monitorada um acúmulo de depósitos de cinzas sobre o equipamento da sessão de transferência de calor convectivo, por meio de medição de perda de carga. Através dos resultados obtidos foi possível identificar precocemente as situações anormais, o modelo possuía robustez a incertezas e ruídos, adaptabilidade a dinâmica do processo, natureza explícita, desenvolvimento e implementação computacional factíveis.

A Tabela 4 apresenta uma relação das principais informações dos estudos descritos anteriormente em relação a utilização de redes neurais nas indústrias de Celulose e Papel. Este estudo do estado da arte foi realizado através de pesquisas nos sites da ScienceDirect, Scielo e o portal de periódicos da Capes, as palavras chaves utilizadas para pesquisa foram “*Tensile Strenght of paper*”, “*Soft sensor in Pulp and Paper Industry*”, “*Neural*

Network in Pulp and Paper Industry”, “Neural Network Tensile Strenght”, “Mathematic models Tensile Strenght”.

Tabela 4 – Estado da Arte da utilização de Redes Neurais na Industria de Celulose e Papel.

Autor	Estrutura da rede	Variável predita	R
Vieira (2003)	<i>Multi-Input-Multi-Output (MIMO)</i>	Resistência a rigidez longitudinal e transversal; Resistência a Delaminação;	MIMO: 0,990 – 0,990; 0,681; 0.638 – 0,586.
	<i>Multi-Input-Single-Output (MISO)</i>	Resistência a Absorção de Ácido Lático e Peróxido de Hidrogênio.	MISO: 0,989 – 0,985; 0,631; 0,920 - 0,940.
Ribeiro (2007)	<i>Multilayer Perceptron</i>	Grau de Cozimento.	0,9276
Stein (2010)	<i>Multilayer Perceptron</i>	Produção do Digestor; Carga Alcalina.	0,88 0,60
Garcia (2010)	-	Marmorização.	0,821
Almeida (2008)	Bayesiana	Eventos de falhas na caldeira	0,924

FONTE: O Autor

A utilização de redes neurais na indústria de celulose e papel, a partir do estado da arte, se mostrou um método amplamente utilizado e com grande aceitação dentro da indústria mas nenhum resultado foi encontrado a partir das palavras chaves para a utilização de redes neurais para a predição da resistência à tração, apesar de encontrar resultados de estudos sobre a resistência à tração.

SHARIFI *et al.* (2016) estudaram polpa destintada, em que foi utilizada para investigar os efeitos das variáveis de branqueamento nas propriedades da polpa e papel. 6 variáveis do processo incluindo peróxido de hidrogênio, hidróxido de sódio, e silicato de sódio, tempo de branqueamento, temperatura do branqueamento e pH final foram otimizados a fim de melhorar os índices de tração. Os resultados mostraram que as variáveis de branqueamento

afetaram significativamente as propriedades ópticas e as propriedades mecânicas do papel, além de melhorar os índices de resistência à tração através de otimização do processo de branqueamento.

CARLSSON e LINDSTROM (2004) examinaram uma abordagem de cisalhamento para a predição da resistência à tração. O estudo demonstrou que a transição da resistência à tração da polpa para o papel requer fibras longas para folhas que possuem fracas ligações fibra-a-fibra ou uma área relativamente baixa de ligações.

LARSON *et al.* (2018) investigaram experimentalmente os efeitos dos comprimentos das fibras e ligações fibra-a-fibra sobre a resistência à tração e tenacidade a fraturas. Eles fizeram folhas em laboratório a partir da polpa Kraft, além de melhorarem a ligação fibra-a-fibra através do enxerto de carboximetil. Os resultados foram comparados com modelos de predição de um atraso de cisalhamento, a resistência à tração e a tenacidade a fraturas em papéis com fracas ligações fibra-a-fibra aumentaram com o comprimento das fibras e os papéis em que foram tratados com carboximetil apresentaram fortes ligações fibra-a-fibra, resultando em altos índices de resistência à tração e tenacidade a fraturas independente do comprimento das fibras.

SCOTT e SHIRT (2001) desenvolveram um modelo de simulação da máquina piloto de produção de papel da Universidade de Miami, este modelo obteve bons resultados para predizer a direção da máquina e a resistência à tração transversal do papel produzido pela máquina piloto. Este modelo inclui os efeitos do refino, nível de carga, nível de amido catiônico e propriedades da fibra. O trabalho também descreve uma maneira de integrar o modelo com uma simulação de balanço de energia e material, além de descrever possíveis aplicações na fabricação do papel como o desenvolvimento de sensores virtuais para monitorar a qualidade do papel, antecipar problemas de qualidade entre outros.

2.5.3 Uso de Redes Neurais Para a Detecção de *Outliers*

Os *outliers* são valores atípicos dos demais valores presentes em um conjunto de dados, normalmente são anomalias, exceções, observações

discordantes, peculiaridades, aberrações ou dados contaminados. É uma observação nos dados em que é altamente improvável de se obter através de um modelo construído com os dados. A detecção de *outliers* é uma parte crucial para o processo de análise de dados e a detecção dos mesmos é um dos problemas mais complexos na análise de dados, e a presença deles se torna muitas vezes enganosa para algoritmos como *clustering*, classificação ou regressão (CHAKRABORTY *et al.*, 2019).

Uma das técnicas de *data mining* para a detecção de anomalias no *dataset* é a utilização de técnicas matemáticas, como redes neurais (PARMAR e PATEL, 2017). Detecção de *Outliers* são derivados de três áreas da computação a quais são: estatística, aprendizado de máquina e redes neurais. As redes neurais podem generalizar bem padrões não vistos e são bons em aprendizado adaptativo, possuem uma tolerância para erros vindo de informação redundante e conseguem gerenciar relações complexas não-linear de *input-output*, podem trabalhar sem informação prévia e são eficientes devido a sua natureza paralela (UPASANI e OM, 2015).

UPASANI e OM (2015), desenvolveram uma rede neural de lógica fuzzy para detecção de outliers, pois consegue raciocinar em situações imprecisas devido a característica da lógica fuzzy de lidar com o incerto. O usuário define um limiar (t) e se o valor associação fuzzy do teste de padrão em uma *hyper-box* for menor que t , então o padrão será declarado como um *outlier* em relação o *hyper-box*. Define-se também um parâmetro p que decide a porcentagem de *hyper-boxes* a serem consideradas para votar no teste de padrão como um outlier.

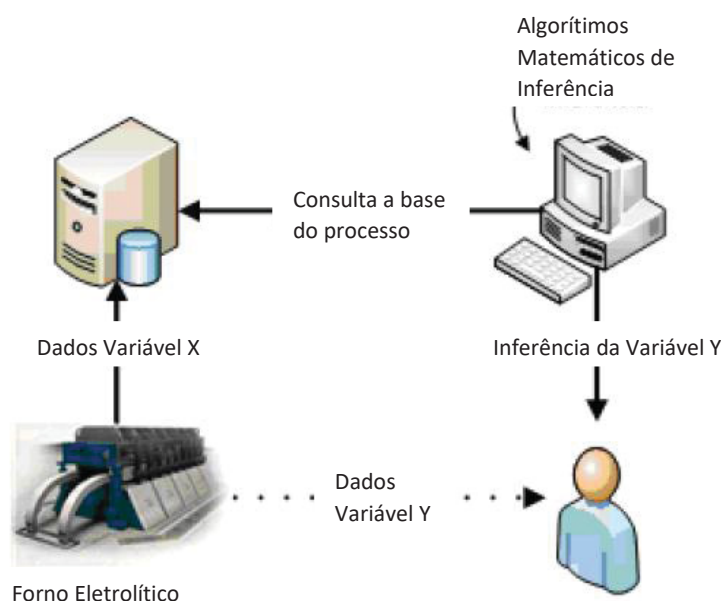
2.6 SENSORES VIRTUAIS

A utilização de sensores virtuais para processos industriais é uma prática antiga, e ela faz com que se remonte o modelo matemático utilizado no processo, tendo como objetivo a predição de propriedades a partir de dados da planta. Os primeiros sensores virtuais desenvolvidos, funcionavam a partir de sistemas de controles indiretos, em que se controlava a variável de interesse a partir de outras variáveis. Ao passar dos anos, e com a evolução da tecnologia e novos estudos sobre a área, estes modelos inferenciais evoluíram com objetivo de estimar uma variável difícil de ser medida, o que levou a grandes esforços para a obtenção de estimadores com maior precisão e confiabilidade (ZANATA, 2005).

Também conhecido como Soft-Sensors, os sensores virtuais permitem a medição de variáveis que não poderiam ser medidas por sensores tradicionais, ou que ainda que os sensores conseguissem medi-las, seria de uma maneira limitada. Eles podem ser considerados um resultado do desenvolvimento e junção das tecnologias de Sensores Inteligentes e das técnicas de Modelagem e Identificação de Sistemas assim, é possível separar os Soft-Sensors em duas partes distintas, sendo elas o conjunto de sensores que medem as variáveis relacionadas a variável desejada e, o modelo do sistema, que é geralmente implementado em software que, através de simulações, é capaz de fornecer a variável desejada (LOTUFO *et al.*, 2008). Um exemplo de estrutura de um sensor virtual pode ser observado na Figura 10.

Por possuírem uma grande área de aplicação, contribuindo para um melhor controle do processo, com medições mais precisas, rápidas e confiáveis e um custo de desenvolvimento, implantação, e manutenção baixos, podendo tanto substituir ou mesmo trabalhar junto com sensores físicos, ajudando no monitoramento, controle de falhas e manutenção preventiva (ZANATA, 2005).

Figura 10 –Diagrama de princípio de sensor virtual.



FONTE: Soares *et al.* (2010).

Os primeiros sensores virtuais foram desenvolvidos por meio de controle indireto, onde a variável de interesse era controlada por meio de variáveis secundárias, a mais comum era a temperatura. Em 1978, Brosilow desenvolveu um dos primeiros sensores, o estimador inferencial (BROSILOW, 1978).

Essa tecnologia tem sido muito aplicada na indústria química, por possuírem processos mais complexos e não lineares, além das variáveis de interesse possuírem um tempo de resposta lento. Em 2010, Soares utilizou sensores virtuais para medir e controlar a temperatura dos fornos de redução de alumínio em que as medições são complexas, pelo fato da natureza corrosiva do material (SOARES *et al.*, 2010).

Os sensores podem enfrentar algumas dificuldades em que a exatidão dos mesmos sofrem uma queda devido a mudanças do estado da planta química, perda de performance da catálise, etc. E se os problemas de degradação do sensor não forem resolvidos, dificulta a identificação de situações anormais. Se o erro de predição de uma variável for maior que o limite aceitável, isso significa que ocorreu uma situação de anomalia e, não

existe uma maneira eficiente para avaliar se a anomalia foi causada por problemas no processo, no analisador ou se o sensor sofreu uma degradação. Um método de resolver o problema de degradação do sensor é a atualização do modelo de regressão com um novo conjunto de dados (LIU, 2007).

O principal método utilizado para o desenvolvimento de um sensor virtual é o método de regressão dos mínimos quadrados parciais (Partial Least Square - PLS). Outros métodos muito utilizados para a construção do sensor são os métodos de regressão por componentes principais (Principal Component Regression - PCR), regressão dos mínimos quadrados parciais não lineares, Redes Neurais Artificiais (RNA) e máquinas de vetores de suporte baseadas em métodos de regressão (KANEKO *et al.*, 2009).

Em 2017, ZHU desenvolveu um sensor virtual para a monitoração do total de efluentes de fósforo (Effluent Total Phosphorus - ETP). Ele utilizou de PLS para selecionar variáveis confiáveis do ETP baseado em dados experimentais. Depois utilizou-se de rede neural de função base radial (Radial Basis Function Neural Network - RBFNN) para identificar a relação entre as variáveis secundárias e ETP, a rede neural facilmente otimizou os parâmetros do modelo para a habilidade de generalização do sensor virtual. Assim desenvolvendo um sistema de monitoramento PLS-RBFNN que apresentou resultados com capacidade de predição melhor que alguns métodos existentes.

Um sensor virtual foi desenvolvido por LIU (2018) para predição de qualidade através de uma estrutura de rede bayesiana. Devido a degradação do modelo, é necessário a construção de um modelo adaptativo capaz de acompanhar as mudanças do processo. O modelo bayesiano demonstrou uma superioridade sobre os outros métodos, especialmente lidando com falta de dados e a habilidade de aprender casualidades. Validou-se o modelo em dados coletados em uma coluna debutanizadora e em uma coluna de absorção de CO₂.

3. MATERIAIS E MÉTODOS

Para a processo de coleta de dados, toda mão de obra técnica, matéria prima e equipamentos necessários foram fornecidos pela KLABIN, sendo eles repassados para a produção deste trabalho.

Coletou-se dados da planta registrados entre 1 de janeiro de 2017 e 1 de novembro de 2018.

O Software utilizado para análise dos dados, construção do modelo matemático e da rede neural, será o software comercial MATLAB cuja licença foi cedida pelo laboratório LACAUT.

3.1. MÉTODOS

3.1.1. Identificação das Propriedades

Primeiramente deve-se selecionar variáveis do processo a serem inferidas (variáveis de saída), a variável a ser estudada é a de resistência à tração do produto. A partir da variável de saída, é possível realizar a identificação das variáveis de entrada do processo que influenciam no mesmo.

3.1.2. Pré-Processamento de Dados

O pré-processamento de dados é o passo mais importante no desenvolvimento de modelos confiáveis. Ele consiste na eliminação de dados não confiáveis que podem ser originados pela baixa velocidade de máquinas, medidas de sensores planas ou sem variação, pontos espúrios. Deve ser verificada as faixas de variação de cada variável de entrada e verificado se as mesmas fazem sentido. Além de eliminar dados relacionados à transição de produtos ou a intervalos em que foram realizados ensaios industriais na máquina. O uso de variáveis correlacionadas do processo, em geral, é problemático no ajuste de modelos empíricos, pois cria problemas numéricos. Nesta etapa realizou-se uma análise de *cluster* para conhecer as características do conjunto de dados.

3.1.3. Análise Estatística de Variáveis

A etapa de seleção de variáveis é de fundamental importância na construção de qualquer modelo, independente de sua tipologia. Ela é responsável pelo sucesso da modelagem ou não. Será utilizado um sistema de baseado em análises multivariáveis, analisando de forma sistemática quais variáveis deverão compor o modelo, para que ele possa representar mais fielmente possível os dados.

Para identificação das propriedades, utilizou-se a correlação de Pearson para uma análise linear e, a correlação de Spearman para análise não linear.

3.1.4. Formulação do Modelo Estatístico

Desenvolveu-se modelos estatísticos baseados em métodos de regressão linear multivariada e regressão por mínimos quadrados parciais multivariados (PLS) a fim de encontrar o modelo que melhor se adéqua ao processo.

3.1.5. Formulação do Modelo Baseado em Redes Neurais

O modelo baseado em redes neurais foi desenvolvido com o auxílio do software MATLAB usando o NNTOOL que é uma GUI (Graphical User Interface) presente no programa.

3.1.6. Formulação do Modelo Estatístico Utilizando PCA No *Dataset*

Utilizou-se da técnica Principal Componente Analysis no conjunto de dados. Desenvolveu-se modelos estatísticos baseados em métodos de regressão linear multivariada e regressão por mínimos quadrados parciais multivariados (PLS) a fim de encontrar o modelo que melhor se adéqua ao processo.

3.1.7. Formulação do Modelo Baseado em Redes Neurais Utilizando PCA No *Dataset*

Utilizou-se da técnica Principal Componente Analysis no conjunto de dados. O modelo baseado em redes neurais foi desenvolvido com o auxílio do software MATLAB.

3.1.8. Formulação do Modelo Estatístico Utilizando RNA Para Detecção De *Outliers* No *Dataset*

Utilizou-se da capacidade da RNA para generalizar e seu alto grau de não-linearidade para detectar outliers no conjunto de dados. Desenvolveu-se modelos estatísticos baseados em métodos de regressão linear multivariada e regressão por mínimos quadrados parciais multivariados (PLS) a fim de encontrar o modelo que melhor se adéqua ao processo.

3.1.9 Formulação do Modelo Baseado em Redes Neurais Utilizando PCA No *Dataset*

Utilizou-se da capacidade da RNA para generalizar e seu alto grau de não-linearidade para detectar outliers no conjunto de dados. O modelo baseado em redes neurais foi desenvolvido com o auxílio do software MATLAB.

3.1.10. Análise de Sensibilidade

A análise de sensibilidade é necessária para conhecer a influência de cada variável de entrada sobre cada variável de saída, assim classificando cada variável de acordo com sua ordem de importância. Caso alguma variável tenha passado pela análise estatística e não possua importância, ela deve ser eliminada do modelo e deve ocorrer um novo treinamento com as variáveis remanescentes.

4. RESULTADOS

4.1. IDENTIFICAÇÃO DAS PROPRIEDADES

A partir do sistema de controle da planta CTMP e conhecendo a variável de interesse (Resistência à Tração) foi possível determinar as variáveis de entrada. Para simplificar a visualização nomeou-se as variáveis entre V1 e V31 e muitas delas podem ser identificadas através do fluxograma do processo presente na revisão.

V1 – Resistência à tração, variável de interesse a ser controlada ($N \times m/g$).

V2 – Sensor referente a medição da balança de cavacos na entrada da planta CTMP (ton/h).

V3 – Velocidade da rosca de descarga do silo de pré-aquecimento (RPM).

V4 – Nível de químicos presente no silo de cavacos (%).

V5 – Velocidade da prensa de cavacos (RPM).

V6 – Torque da prensa de cavacos (KNm).

V7 – Controle de energia da prensa de cavacos (Kwh/t).

V8 – Vazão de químicos para impregnação do cavaco (L/min).

V9 – Velocidade da rosca de descarga do silo de reação (RPM).

V10 – Velocidade do alimentador do refinador 1 (RPM).

V11 – Potência do refinador 1 (MW).

V12 – Velocidade do alimentador do refinador 2 (RPM).

V13 – Potência do refinador 2 (MW).

V14 – Velocidade do filtro de disco (RPM).

V15 – Potência do refinador de rejeitos (MW).

V16 – Velocidade da prensa 1 (RPM).

V17 – Torque da prensa 1 (KNm).

V18 – Torque da prensa 2 (KNm).

V19 – Torque da prensa 3 (KNm).

V20 – Controle de consistência de descarga da bomba 1 (%).

V21 – Controle de consistência de descarga da bomba 2 (%).

V22 – Controle de consistência de descarga para a prensa 1 (%).

V23 – Produção da Impregnação (ton/h).

V24 – Vazão de NaOH para impregnação do cavaco (L/min).

V25 – Pressão de fechamento Refinador 1 (Pa).

V26 – Posição de fechamento do rotor Refinador 1 (mm).

V27 – Pressão de fechamento Refinador 2 (Pa).

V28 – Posição de fechamento do rotor Refinador 2 (mm).

V29 – Produção da planta (ton/dia).

V30 – Densidade aparente da matéria prima (g/cm^3).

V31 – Consumo de energia específico (MW).

4.2. PRÉ-PROCESSAMENTO DE DADOS

Durante o pré-processamento de dados, após eliminar dados provenientes de paradas, falhas de medidas ou baixa confiabilidade, observou-se que o conjunto de dados possuía dois diferentes valores a serem atingidos pela variável de interesse (Resistência à tração), isso se deve a necessidade da KLABIN de se adaptar a demanda do cliente. Após o pré-processamento de dados restaram 1409 conjuntos de dados para a construção do modelo.

Realizou-se uma análise de *cluster* hierárquico nos dados, onde se estudou caso a caso, através de uma associação de 4 clusters, o método de clusterização utilizado foi o de vizinho mais próximo e a medida de intervalo foi a distância euclidiana quadrática. O resultado pode ser analisado na Tabela 5.

Tabela 5 – Número de elementos encontrado em cada *cluster*

Cluster	Nº de Casos	Cluster	Nº de Casos
1	1350	3	1
2	55	4	3

FONTE: O Autor

Observa-se que o primeiro Cluster engloba praticamente todo o conjunto de dados, isso significa que a maior parte dos dados apresentam características similares,

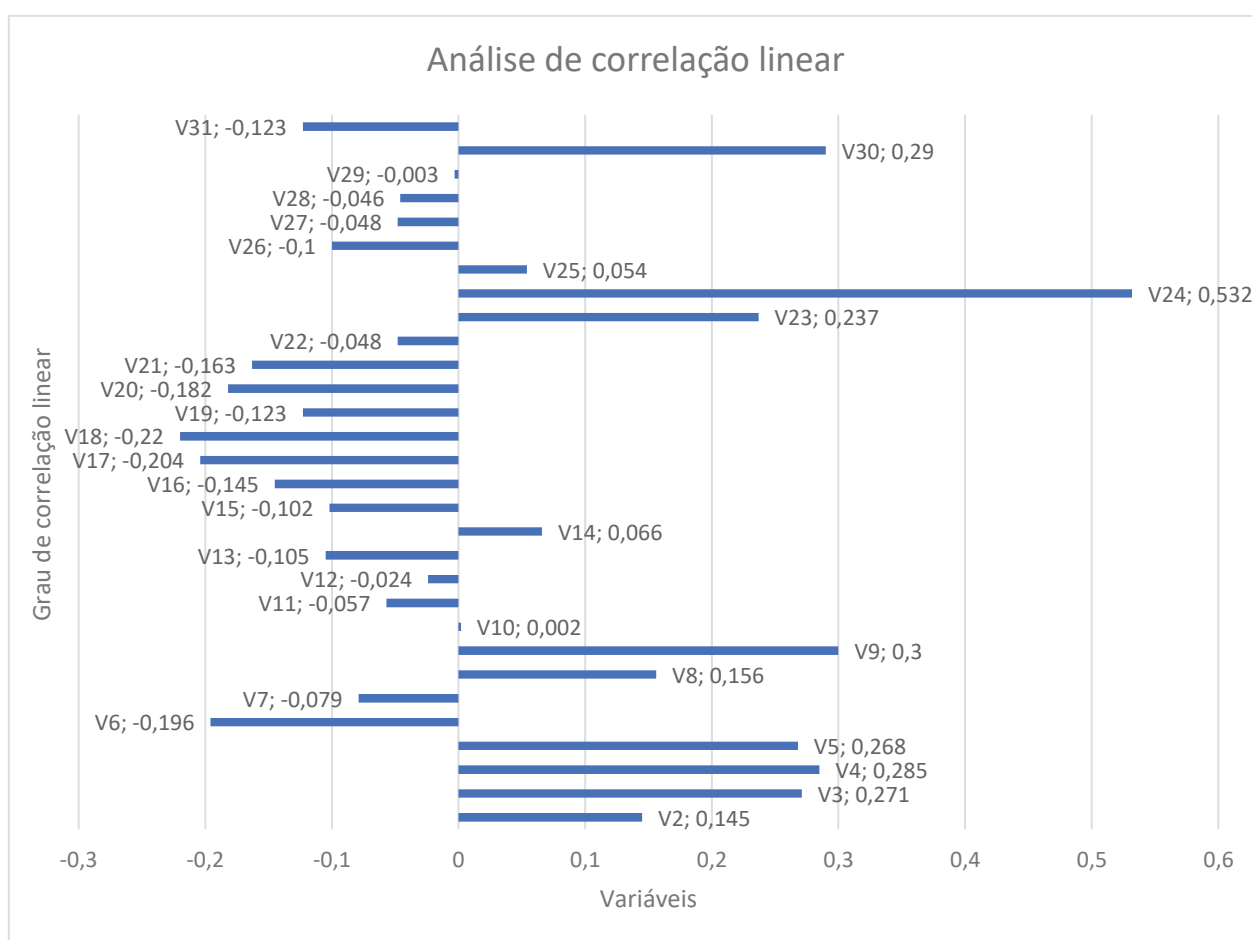
Realizou-se uma normalização nos dados através da Equação 4.1 a fim de minimizar os problemas oriundos do uso de unidades e dispersões distintas entre as variáveis.

$$x_n = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (\text{Equação 4.1})$$

4.3. ANÁLISE ESTATÍSTICA DE VARIÁVEIS

Após o pré-processamento de dados, realizou-se uma análise de correlação linear das variáveis do conjunto de dados através do método de correlação de Pearson, a fim de encontrar as variáveis que possuíam maior influência sobre a variável de interesse. Os resultados dessa análise podem ser observados na Figura 11.

Figura 11 – Análise de correlação linear pelo método de Pearson.

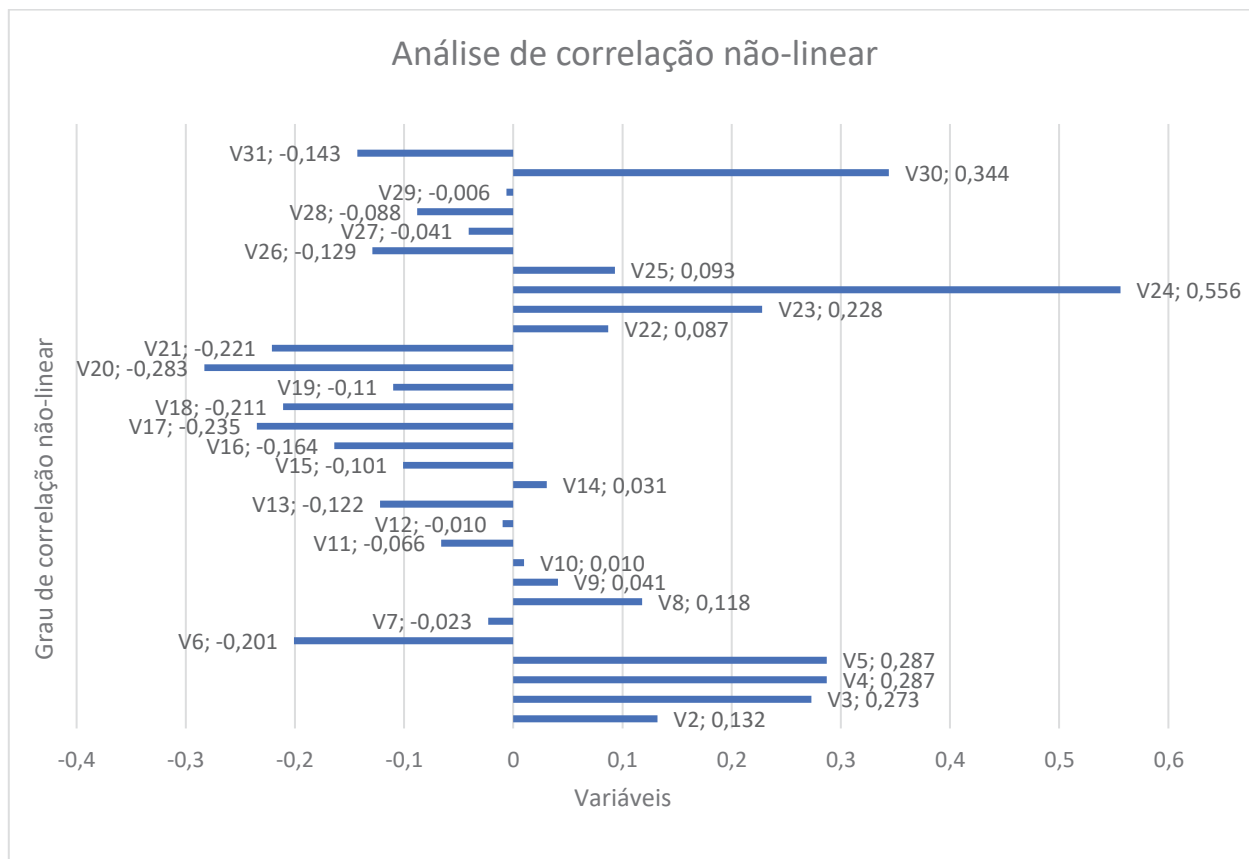


FONTE: O Autor.

É possível observar que a grande maioria das variáveis apresentaram uma correlação fraca, em que apenas a variável de vazão de NaOH para impregnação (V24) apresentou um comportamento moderado com relação a resistência à tração, isso motivou a análise de correlação não-linear das

variáveis através da correlação de Spearman, o resultado pode ser observado na Figura 12.

Figura 12 – Análise de correlação não-linear pelo método de Spearman.



FONTE: O Autor.

Como na correlação linear, a maior parte do conjunto de dados após a análise de correlação não-linear apresentou baixa correlação, apresentando uma correlação moderada apenas nas variáveis de vazão de NaOH para impregnação e a densidade aparente do eucalipto (V30).

Após analisar a correlação das variáveis e, com auxílio do conhecimento técnico dos operadores da KLABIN, selecionou-se um grupo de variáveis para construção do modelo estatístico e para a construção da rede neural. É importante lembrar que correlação não significa causalidade, por esse motivo algumas das variáveis que apresentaram um grau de correlação maior que outras não foram selecionadas para a construção do modelo. Outro fator que influenciou a seleção de variáveis é o fato de que algumas variáveis

que possuem o mesmo grau de correlação carregam informações semelhante, tendo um alto grau de correlação entre si, e que ao selecionarmos as 2 para a construção do modelo em nada iria adicionar na capacidade preditiva, pois possuem a mesma informação, e só aumentaria o trabalho computacional necessária para a construção dos modelos, assim, selecionou-se apenas 1 das variáveis que carregam informação semelhantes entre si. A Tabela 6 apresenta o conjunto de variáveis selecionadas.

Tabela 6 – Variáveis selecionadas para desenvolvimento dos modelos.

Variável	
V2	Balança de cavacos
V3	Velocidade da rosca de descarga do silo de pré-aquecimento
V4	Nível do silo de cavacos
V5	Velocidade da prensa de cavacos
V6	Indicação do torque da prensa de cavacos
V8	Vazão de químicos para impregnação do cavaco
V23	Produção da Impregnação
V24	Vazão de NaOH para impregnação
V30	Densidade aparente da matéria prima
V31	Consumo de energia específico

FONTE: O Autor

4.4. FORMULAÇÃO DO MODELO ESTATÍSTICO

Foi realizado uma regressão linear multivariada a fim de desenvolver um modelo capaz de prever a variável de interesse, os parâmetros do modelo foram descritos da Tabela 7.

Tabela 7 – Parâmetros do modelo linear multivariado.

Variável		Parâmetro
V2	Balança de cavacos	-0,012
V3	Velocidade da rosca de descarga do silo de pré-aquecimento	0,028
V4	Nível do silo de cavacos	0,096
V5	Velocidade da prensa de cavacos	-0,177
V6	Indicação do torque da prensa de cavacos	-0,074
V8	Vazão de químicos para impregnação do cavaco	0,212
V23	Produção da Impregnação	-0,06
V24	Vazão de NaOH para impregnação	0,456
V30	Densidade aparente da matéria prima	0,020
V31	Consumo de energia específico	-0,037
Constante		0,320

FONTE: O Autor

A partir dos coeficientes é possível construir a Equação 4.2, que descreve o modelo linear multivariado.

$$y = -0,012x_1 + 0,028x_2 + 0,096x_3 - 0,177x_4 - 0,074x_5 + 0,212x_6 - 0,06x_7 + 0,456x_8 - 0,020x_9 - 0,037x_{10} + 0,320 \quad (\text{Equação 4.2})$$

Foi feita uma análise deste modelo e os resultados estão presentes na Tabela 8.

Tabela 8 – Análise do modelo de regressão linear múltipla.

R	R²	R² ajustado	Erro padrão de estimativa
0,598	0,358	0,353	0,133

FONTE: O Autor

A partir da análise realizada no modelo linear é possível observar que o mesmo é insuficiente para a predição da variável de interesse, visto que o R² resultou em um valor muito abaixo de um valor aceitável.

Devido os resultados negativos do modelo de regressão linear, desenvolveu-se um modelo pelo método de regressão por mínimos quadrados parciais (PLS). Os parâmetros do modelo estão presentes na Tabela 9.

Tabela 9 – Parâmetros do modelo PLS.

Variável	Parâmetro
V2 Balança de cavacos	-0,014
V3 Velocidade da rosca de descarga do silo de pré-aquecimento	0,021
V4 Nível do silo de cavacos	0,098
V5 Velocidade da prensa de cavacos	-0,164
V6 Indicação do torque da prensa de cavacos	-0,079
V8 Vazão de químicos para impregnação do cavaco	0,208
V23 Produção da Impregnação	-0,068
V24 Vazão de NaOH para impregnação	0,457
V30 Densidade aparente da matéria prima	0,012
V31 Consumo de energia específico	0,04
Constante	0,332

FONTE: O Autor

A partir dos coeficientes é possível construir a Equação 4.3, que descreve o modelo não-linear multivariado.

$$y = -0,014x_1 + 0,021x_2 + 0,098x_3 - 0,164x_4 - 0,079x_5 + 0,208x_6 - 0,068x_7 + 0,457x_8 - 0,012x_9 + 0,04x_{10} + 0,332 \quad (\text{Equação 4.3})$$

Foi feita uma análise deste modelo e os resultados estão presentes na Tabela 10.

Tabela 10 – Análise do modelo PLS.

Fator latente	Variância acumulada de X	Variância de x	Variância de Y	Variância acumulada de Y (R ²)	R ² ajustado
1	0,378	0,378	0,223	0,223	0,222
2	0,116	0,494	0,093	0,316	0,315
3	0,051	0,545	0,035	0,351	0,350
4	0,073	0,618	0,006	0,357	0,355
5	0,07	0,687	0,001	0,358	0,355

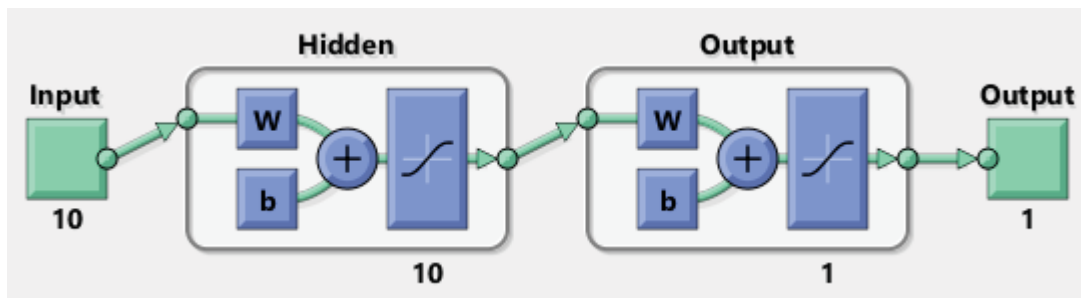
FONTE: O Autor

Novamente o modelo desenvolvido apresentou baixa eficácia na predição da variável de interesse, apresentando um valor baixo do R². Um dos possíveis motivos para esse baixo valor de R² é a provável falta de alguma variável não registrada no conjunto de dados disponíveis, visto que uma das principais variáveis responsáveis pela variação da resistência à tração são variáveis da matéria prima, no qual temos apenas o registro da densidade aparente do cavaco e, um dos principais responsáveis pelo bom desempenho de um modelo matemático se passa pela qualidade do conjunto de dados disponíveis para seu desenvolvimento.

4.5. FORMULAÇÃO DO MODELO BASEADO EM REDES NEURAIS

Com a ajuda do MATLAB, desenvolveu-se uma RNA com função resposta sigmoidal e 1 camada oculta contendo 10 neurônios. A estrutura da rede pode ser observada na Figura 13.

Figura 13 – Estrutura da Rede Neural Artificial desenvolvida.

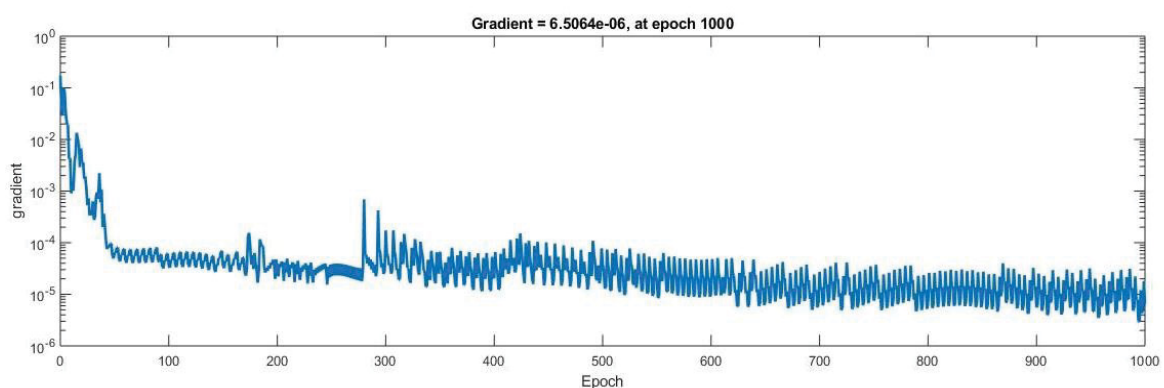


FONTE: O Autor.

Essa estrutura foi escolhida pois apresentou melhor desempenho comparada as outras funções de ativação e variação do número de neurônios na camada oculta. O conjunto de dados foi dividido aleatoriamente, sendo 60% dos pontos presentes no *dataset* utilizados para treinamento da rede, 20% dos pontos utilizados para a validação da rede e 20% dos pontos sendo separados como dados virgens para realização dos testes da rede.

O estado do treinamento pode ser observado na Figura 14.

Figura 14 – Estado do treinamento da Rede Neural Artificial.



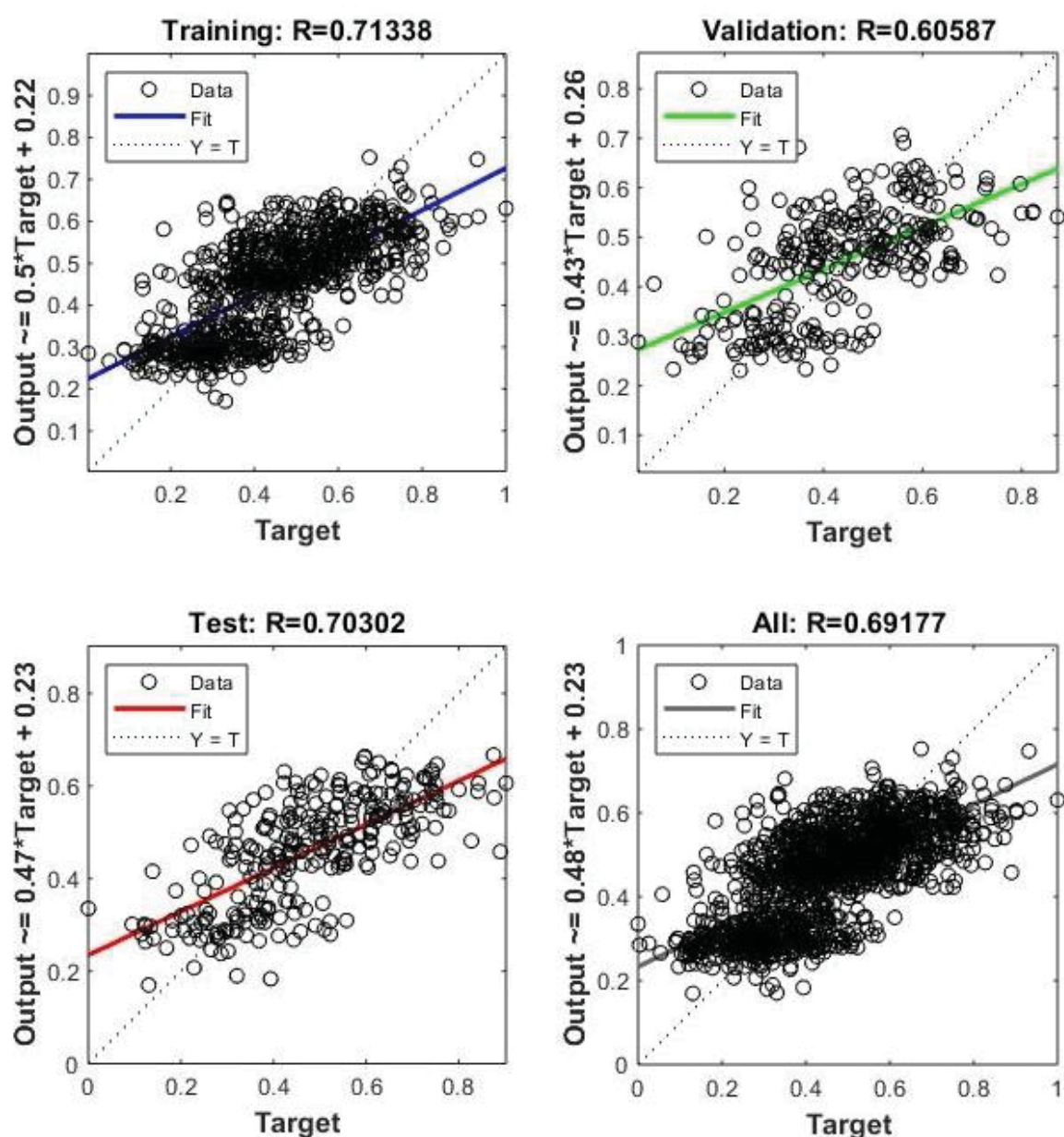
FONTE: O Autor.

É possível observar através da Figura 14 que há uma estabilidade no processo de treinamento, isso indica que o tempo foi suficiente para completar

o treinamento da rede e que, aumentar o tempo de treinamento não iria acarretar em uma melhora significativa no modelo, podendo inclusive gerar problemas de *overfit*, em que o modelo fica viciado nos dados de treinamento o que faz com que a resposta do modelo a novas entradas não seja um valor satisfatório.

Os resultados obtidos na rede neural estão presentes na Figura 15.

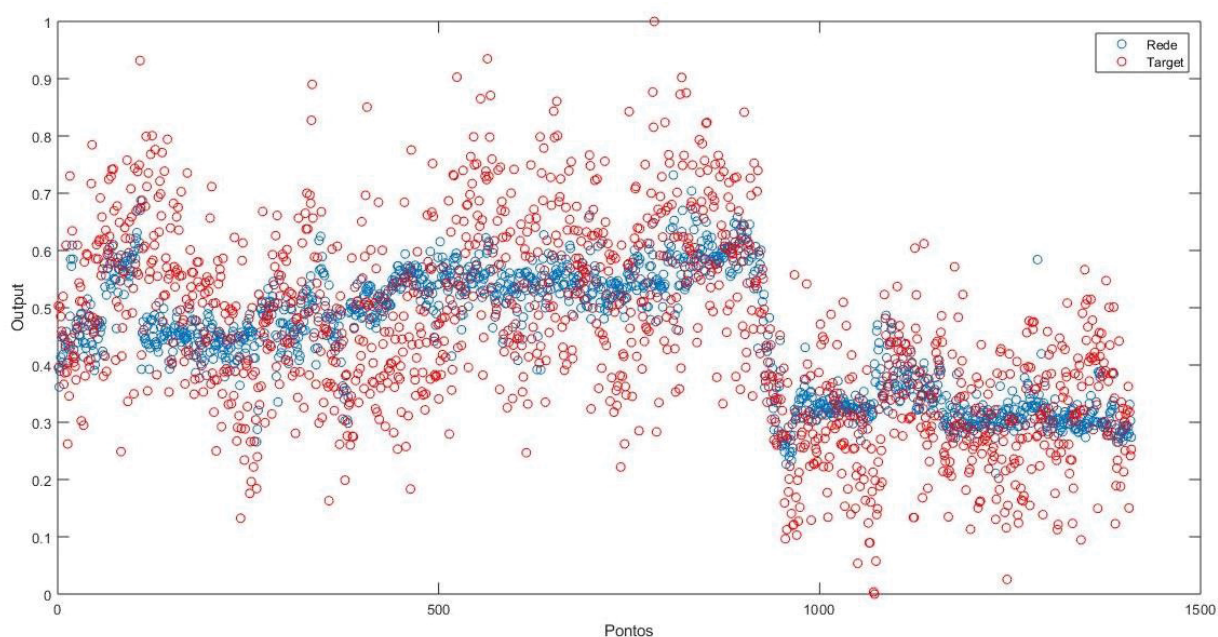
Figura 15 – Regressão do modelo baseado em Redes Neurais.



FONTE: O Autor.

A Figura 15 é composta por 4 gráficos, sendo eles a plotagem da comparação entre o resultado esperado e resultado da rede para os dados de treinamento, validação, teste e um conjunto contendo todos os dados. Ela mostra que a regressão do modelo treinado teve como R um valor de 0,71338, o que mostra que o a rede desenvolvida não possui um desempenho satisfatório, como pode ser observada no teste da rede e na validação da rede, que apresentaram valores baixos de R e próximos a rede treinada. É importante lembrar que, como os dados foram divididos aleatoriamente, a rede irá apresentar valores diferentes de R para os conjuntos de treinamento, validação e teste, assim, uma análise global se torna o melhor método para avaliar a rede, tendo como R um valor de 0,69177. A Figura 16 mostra a comparação dos resultados obtidos com os resultados esperados.

Figura 16 – Comparação dos resultados da rede e dos resultados reais.

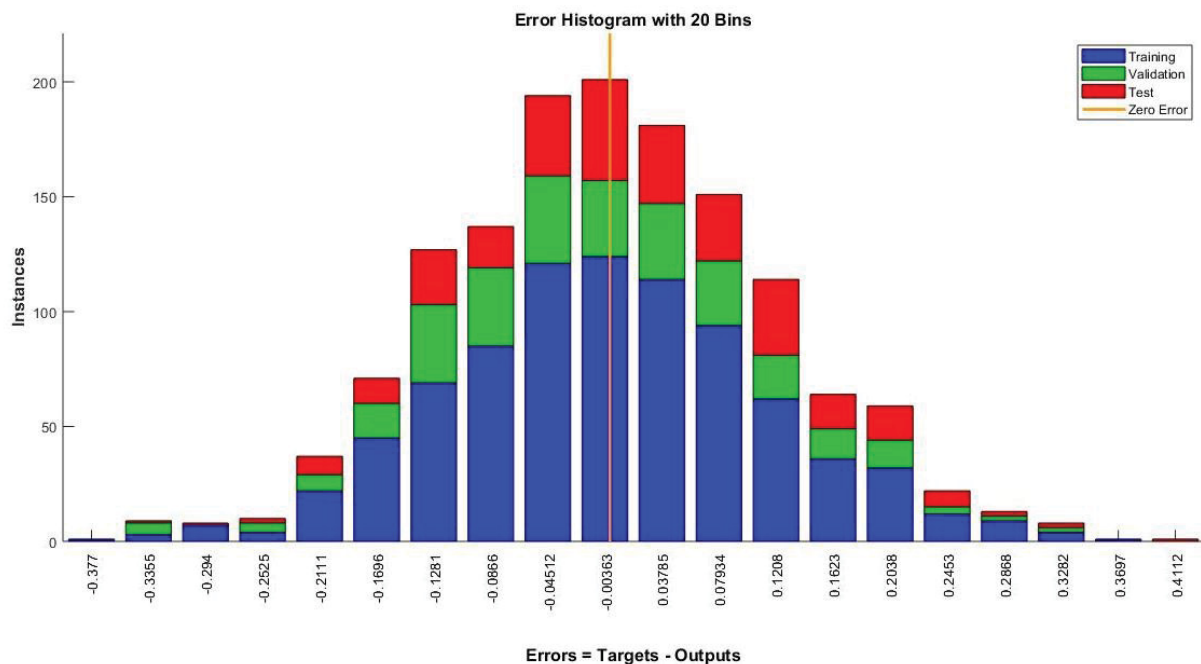


FONTE: O Autor.

É possível observar através da Figura 16 que apesar do baixo desempenho apresentado pela rede, ela conseguiu capturar a tendência do sistema a ser simulado.

. A Figura 17 mostra o erro absoluto da rede.

Figura 17 – Histograma de erro da Rede Neural Artificial.



FONTE: O Autor.

Observa-se uma grande quantidade de pontos com uma faixa de erro absoluto entre 0,1 e 0,2 seja positivamente ou negativamente, apesar da grande maioria estar faixa de erro aceitável pela empresa que é 0,12. Como a rede teve um tempo de treinamento satisfatório e através da análise dos resultados, percebe-se um erro sistemático do conjunto de dados utilizado para desenvolver a rede, gerando falta de confiabilidade dos dados a serem trabalhados. Outra possibilidade pode ser a falta de medição de alguma variável de vital importância na predição da variável de interesse.

4.6.FORMULAÇÃO DO MODELO ESTATÍSTICO UTILIZANDO PCA NO DATASET

Como os modelos desenvolvidos apresentaram um desempenho não satisfatório, e analisando os resultados obtidos pelos modelos, detectou-se a necessidade de utilizar uma nova abordagem sobre o *dataset*. Uma das técnicas aplicadas sobre o *dataset* foi a de Análise dos Componentes Principais (PCA), pois esta técnica permitiria reduzir a dimensionalidade do

modelo, o que poderia estar acarretando erro, além de garantir que nenhuma variável irá apresentar nenhum tipo de colinearidade.

Os resultados da porcentagem em que cada nova variável explica o conjunto de dados pode ser observada na Tabela 11.

Tabela 11 – Porcentagem explicativa de cada variável

Variável	Autovalor	%	Variável	Autovalor	%
PC1	3,938	42,3700	PC6	0,787	6,8715
PC2	1,161	12,0268	PC7	0,466	3,9022
PC3	1,103	10,1687	PC8	0,323	3,2298
PC4	1,011	9,3990	PC9	0,227	2,4931
PC5	0,866	8,2137	PC10	0,168	1,3253

FONTE: O Autor

A partir destes resultados, com base no critério de Kaiser, selecionou-se as quatro principais variáveis novas obtidas para formar um novo conjunto de dados para desenvolver o sensor. Desenvolveu-se um modelo de regressão linear multivariado, os parâmetros estão presentes na Tabela 12.

Tabela 12 – Parâmetros do modelo linear multivariado (PCA)

Variável	Parâmetro	Variável	Parâmetro
PC1	0,207	PC3	-0,041
PC2	-0,222	PC4	0,082
Constante	0,453		

FONTE: O Autor

A partir dos coeficientes construiu-se a Equação 4, que descreve o modelo linear multivariado baseado no *dataset* aplicado PCA.

$$y = 0,207x_1 - 0,222x_2 - 0,041x_3 + 0,082x_4 + 0,453 \quad (\text{Equação 4})$$

Foi feita uma análise deste modelo e os resultados estão presentes na Tabela 13.

Tabela 13 – Análise do modelo de regressão linear múltipla (PCA).

R	R²	R² ajustado	Erro padrão de estimativa
0,509	0,259	0,257	0,143

FONTE: O Autor

A partir da análise realizada no modelo de regressão linear (PCA), observa-se novamente que o modelo apresentou um resultado insuficiente na predição da variável de interesse, visto que o R² resultou em um valor muito abaixo e menor do que o modelo linear anterior.

Devido a insuficiência do modelo anterior, construiu-se um modelo de regressão por mínimos quadrados parciais utilizando o novo *dataset*. Os parâmetros do modelo estão presentes na Tabela 14.

Tabela 14 – Parâmetros do modelo PLS (PCA)

Variável	Parâmetro	Variável	Parâmetro
PC1	0,207	PC3	-0,041
PC2	-0,222	PC4	0,082
Constante	0,453		

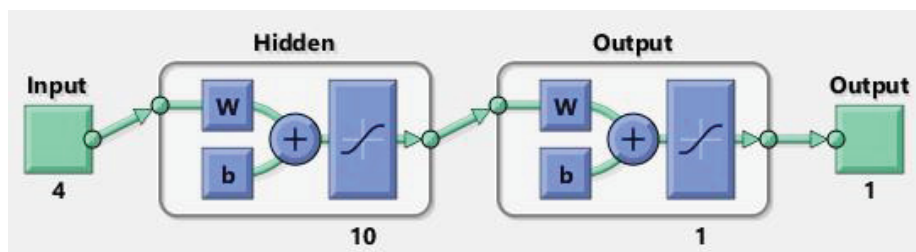
FONTE: O Autor

Observa-se que os parâmetros encontrados são idênticos aos parâmetros do modelo anterior, gerando a mesma equação. Isto se deve ao fato de que a regressão por mínimos quadrados tem como objetivo a eliminação de colinearidade entre as variáveis independentes e, como foi aplicado a técnica de *Principal Component Analysis* em nosso conjunto de dados, garantindo que não exista nenhuma colinearidade entre as variáveis, ao fazermos a regressão encontrou-se o mesmo modelo. O que implica que o modelo anterior já está ajustado.

4.7.FORMULAÇÃO DO MODELO BASEADO EM REDES NEURAIS UTILIZANDO PCA NO *DATASET*

Como o modelo anterior apresentou resultado insatisfatório, desenvolveu-se uma rede neural artificial utilizando o *dataset* em que se aplicou a técnica de *Principal Component Analysis*. A rede foi estruturada contendo as 4 variáveis independentes de entrada e 1 variável objetivo, com função resposta sigmoideal e 1 camada oculta contendo 10 neurônios. A estrutura da rede pode ser observada na Figura 18.

Figura 18 – Estrutura da Rede Neural Artificial (PCA).

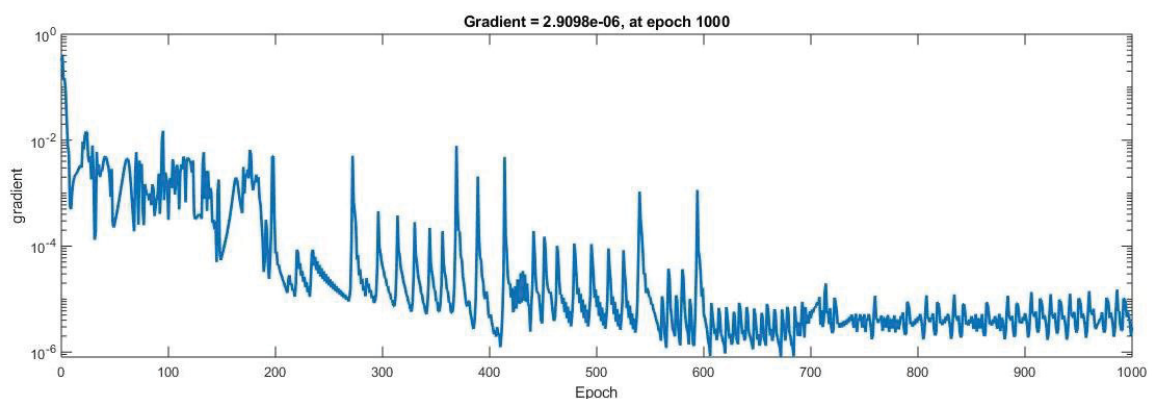


FONTE: O Autor.

Essa estrutura apresentou melhor desempenho comparada as outras funções de ativação e variação do número de neurônios na camada oculta. O conjunto de dados foi dividido aleatoriamente, sendo 60% dos pontos presentes no *dataset* utilizados para treinamento da rede, 20% dos pontos utilizados para a validação da rede e 20% dos pontos sendo separados como dados virgens para realização dos testes da rede.

O estado do treinamento da rede pode ser observado na Figura 19.

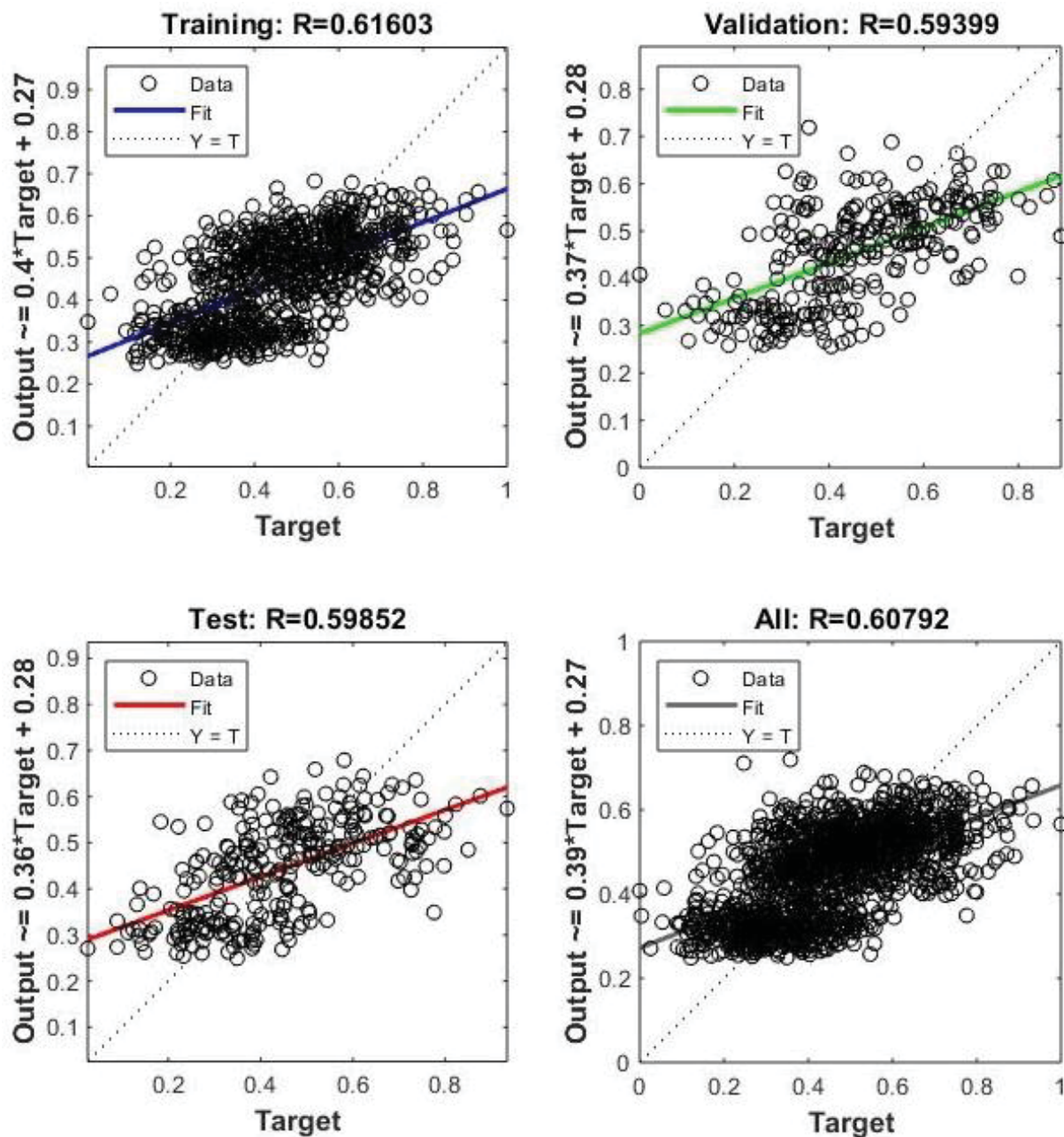
Figura 19 – Estado do treinamento da Rede Neural Artificial (PCA).



FONTE: O Autor.

É possível observar através da Figura 19 que a rede atinge uma estabilidade no processo de treinamento, e que, aumentar o tempo de treinamento não iria gerar uma melhora significativa no modelo. Após o treinamento da rede, realizou-se uma análise sobre o desempenho da mesma, o resultado pode ser observado na Figura 20.

Figura 20 – Regressão do modelo baseado em Redes Neurais (PCA).

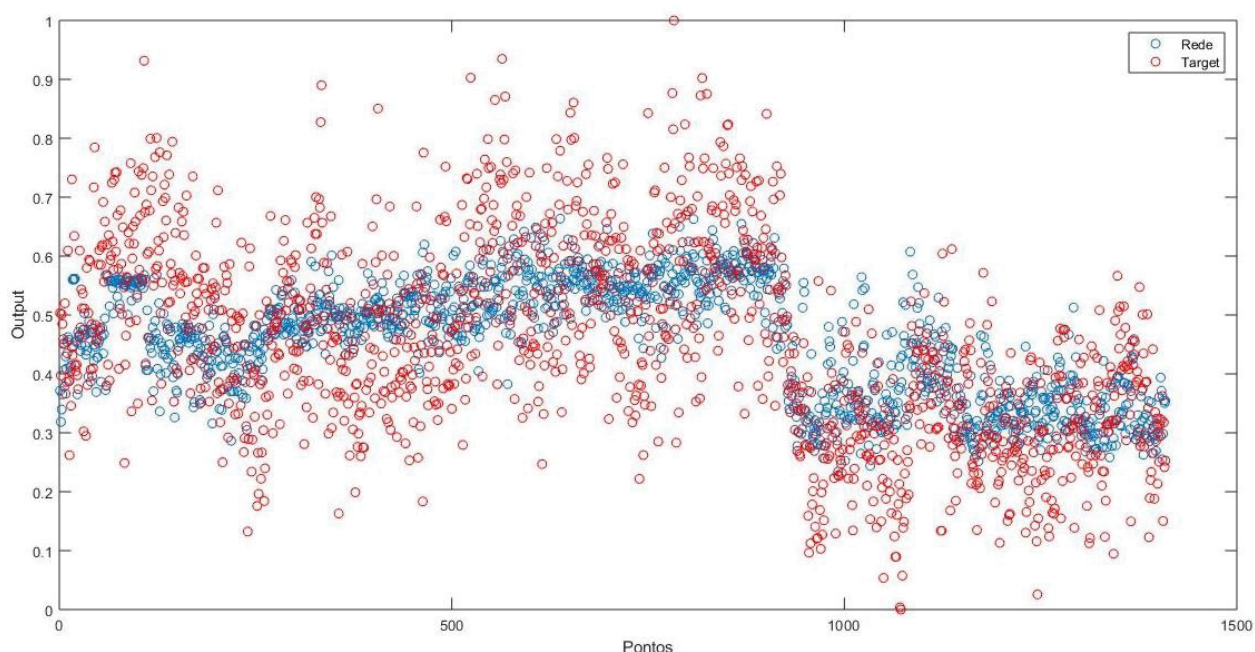


FONTE: O Autor.

A Figura 20 mostra que a regressão do modelo treinado teve como R um valor de 0,61603, o que mostra que o a rede desenvolvida não possui um desempenho satisfatório e inferior a rede treinada anteriormente, como pode ser observada no teste da rede e na validação da rede, que apresentaram valores menores de R comparada a rede anterior.

Realizou-se uma análise da saída da rede para a detecção de possíveis erros. A Figura 21 mostra a comparação dos resultados obtidos com os resultados esperados.

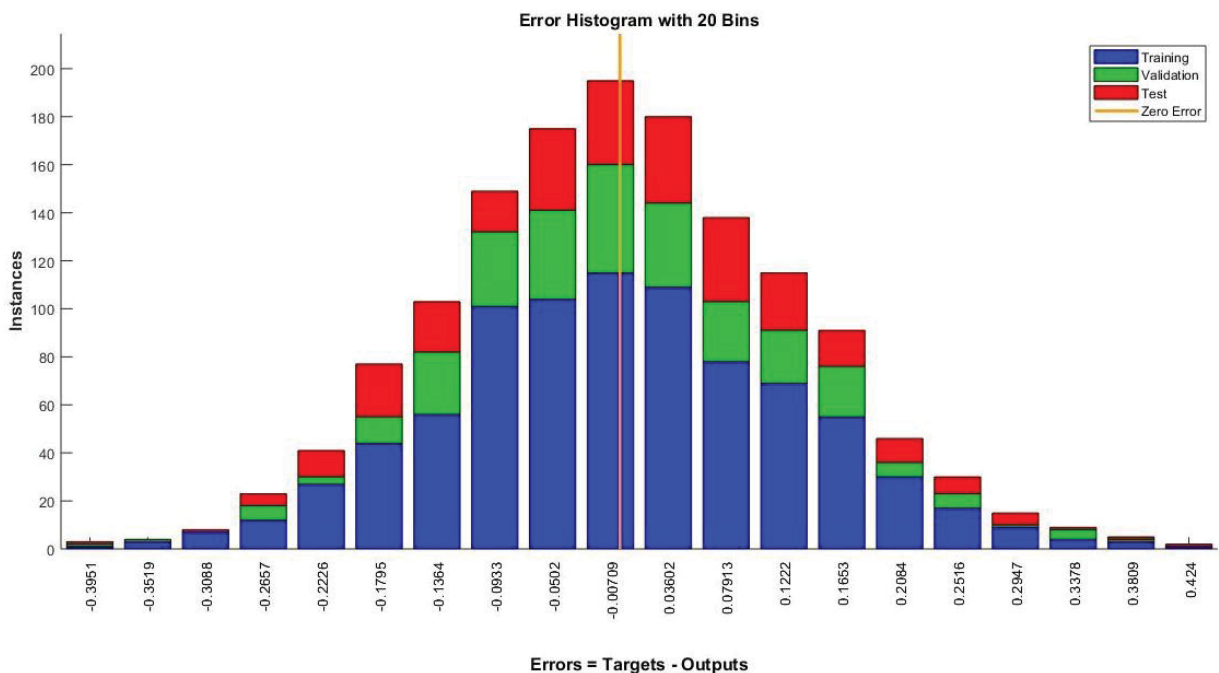
Figura 21 – Comparação entre a RNA e os resultados reais (PCA).



FONTE: O Autor.

É possível observar novamente através da Figura 21 que apesar do baixo desempenho apresentado pela rede, ela conseguiu capturar a tendência do sistema a ser simulado, o que indica um problema de pontos em nosso conjunto de dados influenciando sobre o treinamento da rede, fazendo-se necessário uma nova abordagem sobre nosso banco de dados, a fim de encontrar pontos que estão influenciando diretamente no resultado do treinamento da rede. A Figura 22 tem como objetivo fazer uma análise mais profunda dos erros da rede desenvolvida com o *dataset* aplicado PCA.

Figura 22 – Histograma de erro da Rede Neural Artificial (PCA).



FONTE: O Autor.

Observa-se um aumento da quantidade de pontos com uma faixa de erro absoluto entre 0,1 e 0,2 seja positivamente ou negativamente, comparado à rede treinada anteriormente, junto com uma diminuição dos pontos em que o erro absoluto é próximo a 0. Percebe-se que mesmo aplicando PCA o erro sistemático do conjunto de dados utilizado para desenvolver a rede persistiu, mostrando que o PCA foi um método ineficaz para a resolução do problema deste conjunto de dados.

4.8. FORMULAÇÃO DO MODELO ESTATÍSTICO UTILIZANDO RNA PARA DETECÇÃO DE *OUTLIERS* NO *DATASET*

Por se tratar de um sistema complexo que possui muitas dimensões, é provável que o pré-tratamento para remoção de pontos não detectou prováveis outliers que por passarem despercebidos, consequentemente, afetou o desenvolvimento do modelo e, para solucionar este problema, uma nova abordagem sobre o conjunto de dados se fez necessária. Devido à forte característica não-linear das redes neurais e sua capacidade de aprendizagem, utilizou-se dos resultados encontrados pela rede neural para detectar e remover os outliers e selecionar um novo conjunto de dados que

fossem capazes de formular um modelo com desempenho satisfatório para o sensor virtual. A partir da eliminação desses pontos, formou-se um dataset mais enxuto, possibilitando a construção de novos modelos a partir deste dataset. Desenvolveu-se um modelo baseado em regressão linear multivariada, os parâmetros do modelo foram descritos na Tabela 15.

Tabela 15 – Parâmetros do modelo linear multivariado (RNA).

Variável	Parâmetro
V2 Balança de cavacos	0,027
V3 Velocidade da rosca de descarga do silo de pré-aquecimento	-0,007
V4 Nível do silo de cavacos	0,058
V5 Velocidade da prensa de cavacos	-0,353
V6 Indicação do torque da prensa de cavacos	-0,125
V8 Vazão de químicos para impregnação do cavaco	0,333
V23 Produção da Impregnação	-0,020
V24 Vazão de NaOH para impregnação	0,780
V30 Densidade aparente da matéria prima	0,084
V31 Consumo de energia específico	-0,015
Constante	0,271

FONTE: O Autor

A partir dos coeficientes é possível construir a Equação 4.5, que descreve o modelo linear multivariado.

$$y = 0,027x_1 - 0,007x_2 + 0,058x_3 - 0,353x_4 - 0,125x_5 + 0,333x_6 - 0,02x_7 + 0,780x_8 + 0,084x_9 + -0,015x_{10} + 0,271 \quad (\text{Equação 4.5})$$

Realizou-se uma análise deste modelo e os resultados estão presentes na Tabela 16.

Tabela 16 – Análise do modelo de regressão linear múltivariado (RNA).

R	R²	R² ajustado	Erro padrão de estimativa
0,834	0,696	0,692	0,113

FONTE: O Autor

Observa-se que com a formação do novo conjunto de dados, foi possível desenvolver um modelo com melhor resultado comparado aos modelos desenvolvidos pelo conjunto de dados iniciais e os modelos desenvolvidos pelo conjunto de dados após aplicarmos PCA. Isto mostra que o conjunto de dados inicial apresentava problemas de medição, sendo eles provenientes de possíveis descalibragem de equipamentos de inferência presentes no processo ou possíveis variações laboratoriais devido a mudança de precisão dos técnicos responsáveis pela medição.

Para fins de comparação de desempenho, construiu-se um modelo de regressão por mínimos quadrados parciais utilizando o novo *dataset*. Os parâmetros do modelo estão presentes na Tabela 17.

Tabela 17 – Parâmetros do modelo PLS (RNA).

Variável	Parâmetro
V2 Balança de cavacos	0,025
V3 Velocidade da rosca de descarga do silo de pré-aquecimento	-0,002
V4 Nível do silo de cavacos	0,060
V5 Velocidade da prensa de cavacos	-0,340
V6 Indicação do torque da prensa de cavacos	-0,137
V8 Vazão de químicos para impregnação do cavaco	0,332
V23 Produção da Impregnação	-0,038
V24 Vazão de NaOH para impregnação	0,781
V30 Densidade aparente da matéria prima	0,072
V31 Consumo de energia específico	-0,019
Constante	0,290

FONTE: O Autor

A partir dos coeficientes é possível construir a Equação 4.6, que descreve o modelo não-linear multivariado.

$$y = 0,025x_1 - 0,002x_2 + 0,060x_3 - 0,340x_4 - 0,137x_5 + 0,332x_6 - 0,038x_7 + 0,781x_8 - 0,072x_9 - 0,019x_{10} + 0,290$$

(Equação 4.6)

Realizou-se uma análise deste modelo e os resultados estão presentes na Tabela 18.

Tabela 18 – Análise do modelo PLS (RNA).

Fator latente	Variância acumulada de X	Variância de x	Variância de Y	Variância acumulada de Y (R ²)	R ² ajustado
1	0,388	0,388	0,430	0,430	0,430
2	0,121	0,509	0,177	0,608	0,607
3	0,054	0,563	0,070	0,677	0,676
4	0,061	0,624	0,017	0,695	0,693
5	0,061	0,685	0,001	0,696	0,694

FONTE: O Autor

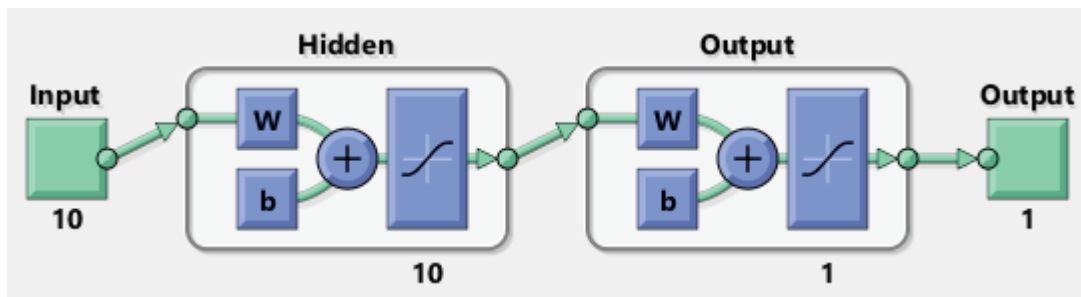
O modelo de regressão por mínimos quadrados parciais desenvolvido apresentou um desempenho superior na predição da variável de interesse em comparação aos modelos anteriores, apresentando um valor superior para R² e com boa performance para predição, mostrando novamente a eficiência da rede neural para a correção do problema de conjunto de dados.

4.9.FORMULAÇÃO DO MODELO BASEADO EM REDES NEURAIIS UTILIZANDO RNA PARA DETECÇÃO DE *OUTLIERS* NO *DATASET*

Como os modelos baseados em redes neurais desenvolvidos anteriormente apresentaram um desempenho superior comparado aos modelos estatísticos com os respectivos conjuntos de dados trabalhados, desenvolveu-se uma rede neural a partir do novo conjunto de dados esperando uma rede com capacidade preditiva superior a todos os outros modelos desenvolvidos. A rede foi estruturada contendo as 10 variáveis independentes de entrada e 1 variável objetivo, com função resposta sigmoidal e 1 camada oculta contendo 10 neurônios, uma estrutura

semelhante a primeira rede construída, e que apresentou melhor desempenho. A estrutura da rede pode ser observada na Figura 23.

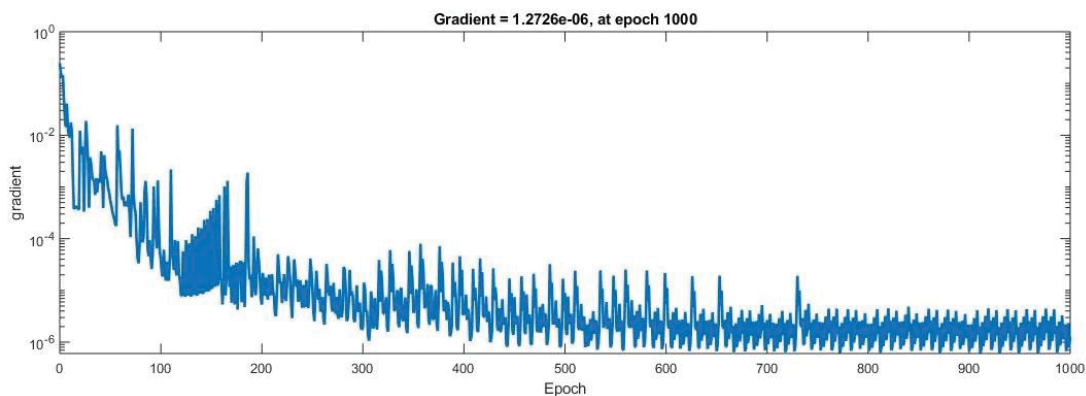
Figura 23 – Estrutura da Rede Neural Artificial (Conjunto de dados RNA).



FONTE: O Autor.

O conjunto de dados foi dividido aleatoriamente, sendo 60% dos pontos presentes no *dataset* utilizados para treinamento da rede, 20% dos pontos utilizados para a validação da rede e 20% dos pontos para realização dos testes da rede. O estado do treinamento da rede pode ser observado na Figura 24.

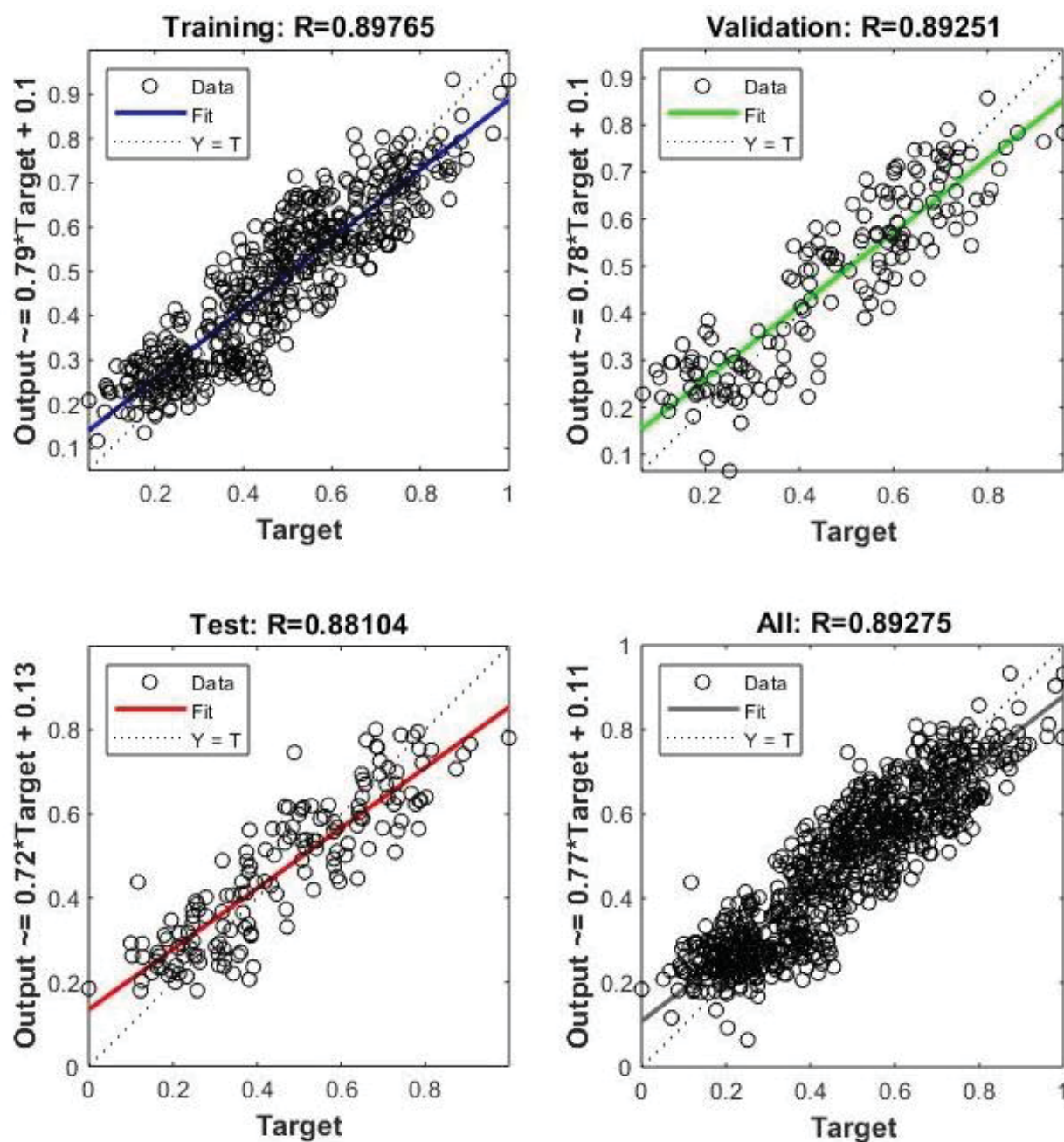
Figura 24 – Estado do treinamento da Rede Neural Artificial (Conjunto de dados RNA).



FONTE: O Autor.

Observar-se através da Figura 24 que a rede atinge uma estabilidade no processo de treinamento, e que, aumentar o tempo de treinamento não iria gerar uma melhora significativa no modelo. Após o treinamento da rede, realizou-se uma análise sobre o desempenho da mesma, o resultado pode ser observado na Figura 25.

**Figura 25 – Regressão do modelo baseado em Redes Neurais
(Conjunto de dados RNA).**

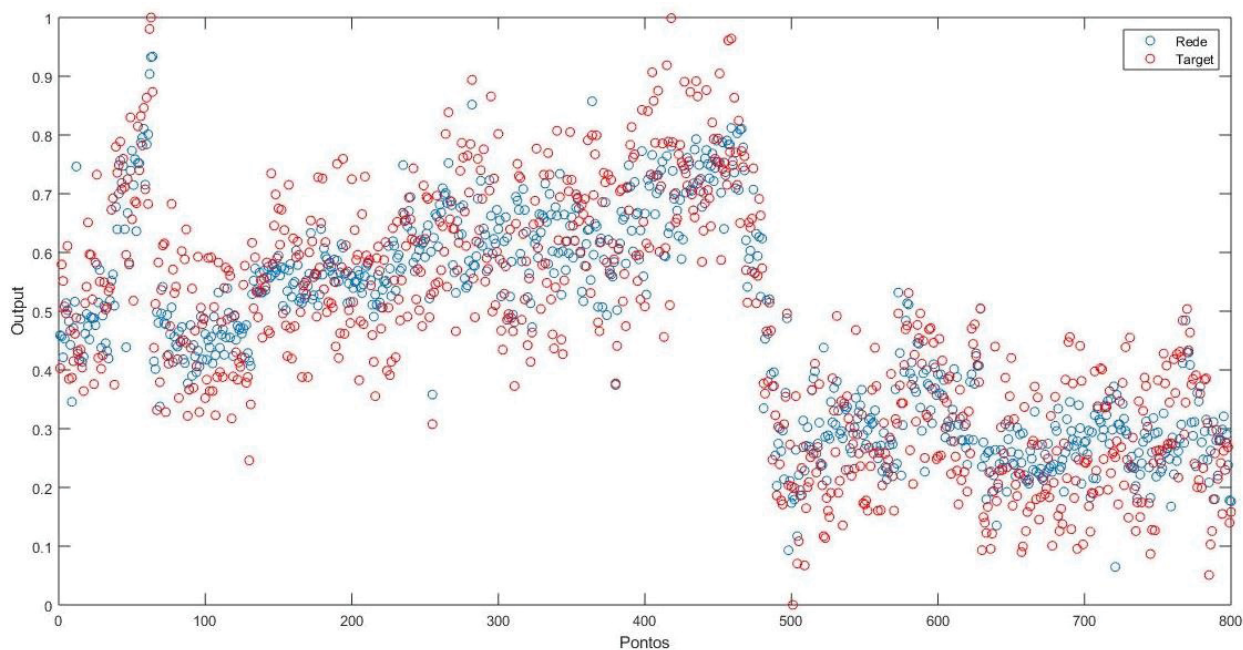


FONTE: O Autor.

A Figura 25 mostra que a regressão do modelo treinado teve como R um valor de 0,89765, o que mostra que a rede teve um desempenho muito superior aos outros modelos desenvolvidos apresentando um resultado com valores do teste da rede e sua validação próximos ao encontrado pela rede

treinada, sendo estes resultados satisfatório para o desenvolvimento de um sensor virtual levando em consideração que tem como objetivo prever uma variável de um processo químico industrial. A Figura 26 mostra a comparação dos resultados obtidos com os resultados esperados.

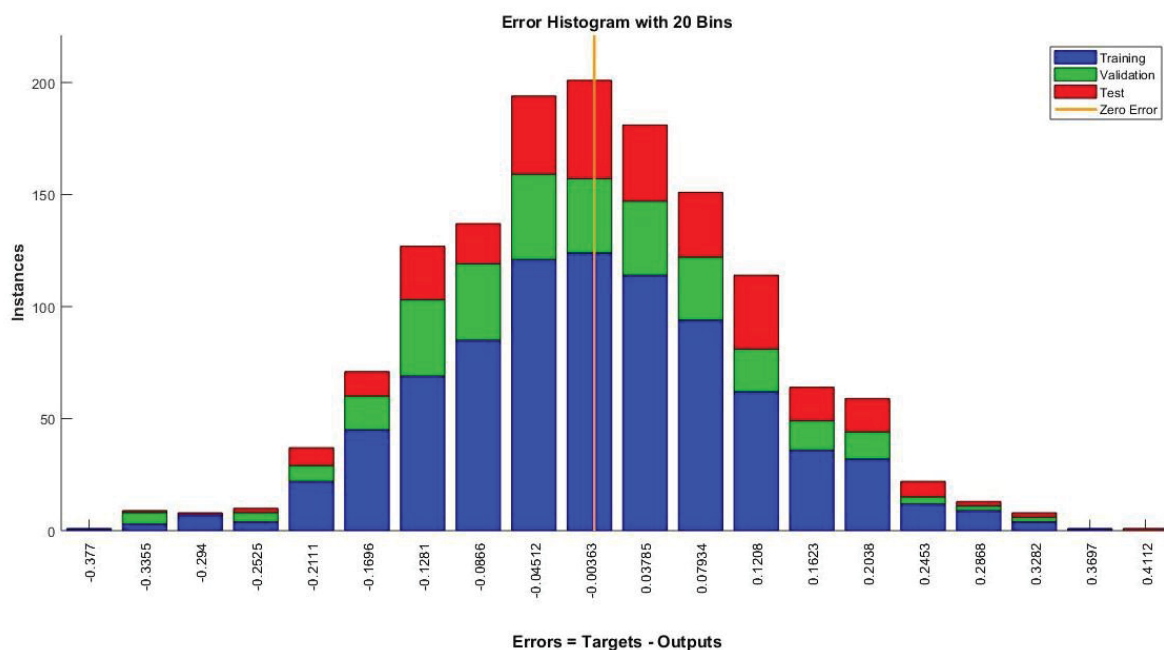
Figura 26 – Comparação entre a RNA e os resultados reais (Conjunto de dados RNA).



FONTE: O Autor.

É possível observar novamente através da Figura 26 ela conseguiu capturar a tendência do sistema a ser simulado além de diminuir a margem de erro da rede em comparação aos outros modelos desenvolvidos, mostrando ser eficiente para realização de controle da planta CTMP. A Figura 27 tem como objetivo fazer uma análise mais profunda dos erros da rede desenvolvida com o *dataset* aplicado redes neurais para remoção de outliers.

Figura 27 – Histograma de erro da Rede Neural Artificial (Conjunto de dados RNA).



FONTE: O Autor.

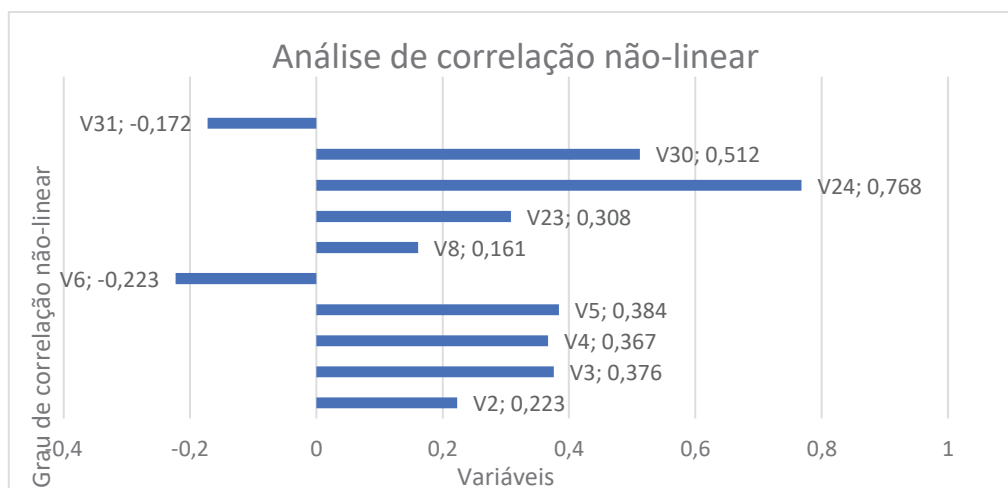
Observa-se que após a eliminação dos outliers, ocorre uma diminuição da quantidade de pontos com uma faixa de erro absoluto entre 0,12 e 0,2 seja positivamente ou negativamente, comparado as outras redes treinadas, e um aumento da porcentagem do número de pontos em que o erro absoluto está dentro da faixa de erro aceitável pela a empresa, mostrando a eficácia do novo processo de remoção de outliers. Percebe-se que ao remover esses outliers a rede apresentou resultados satisfatórios e superiores as outras redes treinadas, conseguindo sobrepor o erro sistemático apresentado com através do seu grande poder de generalização e alto grau de não-linearidade deste método, além de demonstrar que um dos principais problemas ao se trabalhar no desenvolvimento de sensores virtuais passa pela qualidade do conjunto de dados disponível para a construção de modelos matemáticos.

4.10. ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS

Após o desenvolvimento do modelo capaz de prever com eficácia a variável objetivo, realizou-se uma análise de sensibilidade das variáveis de controle do processo a fim de observar qual variável possui maior influência

sobre a resistência à tração durante o processo de produção, para esta análise levou-se em conta os valores encontrados pela correlação de Spearman, os resultados foram apresentados na Figura 28.

Figura 28 – Análise de correlação não-linear pelo método de Spearman no novo dataset.



FONTE: O Autor.

A partir destes dados é possível estabelecer um *rank* para determinar qual a ordem de variáveis que exercem maior influência sobre o controle da resistência à tração. A ordem de influência pode ser observada na Tabela 19.

Tabela 19 – Análise de sensibilidade das variáveis do processo.

Variável	Parâmetro
V24 Vazão de NaOH para impregnação	0,768
V30 Densidade aparente da matéria prima	0,512
V5 Velocidade da prensa de cavacos	0,384
V3 Velocidade da rosca de descarga do silo de pré-aquecimento	0,376
V4 Nível do silo de cavacos	0,367
V23 Produção da Impregnação	0,308
V2 Balança de cavacos	0,223
V8 Vazão de químicos para impregnação de cavacos	0,161
V31 Consumo de energia específico	-0,172
V6 Indicação do torque na prensa de cavacos	-0,223

FONTE: O Autor

Observa-se que a etapa de tratamento químico para amolecer o cavaco possui grande influência na resistência à tração do produto da planta de produção CTMP, muito devido a quantidade de NaOH presente na etapa de impregnação, como indica a sensibilidade da variável objetivo (resistência à tração) em relação a variável V24 (Vazão de NaOH para impregnação). SHARIFI (2016) em seus estudos demonstrou que o NaOH tem influência direta sobre a resistência à tração.

Outra variável que apresentou grande influência sobre a resistência à tração é a V30 (Densidade aparente do Eucalipto da CTMP), que é uma variável que carrega informações sobre a matéria prima sem esta, a única variável que os operadores possuem informação sobre matéria prima e, como vimos na literatura a matéria prima possui grande influência sobre a resistência à tração, pois ela está relacionada ao grau de ligação entre fibras, onde polpas com uma maior quantidade de fibra por grama possuem um maior índice de resistência à tração, devido ao maior contato entre as fibras, o que aumenta as suas ligações (D'ALMEIDA, 1998).

Outro ponto importante para ressaltar são as variáveis V6 (Indicação do torque da prensa de cavacos MSD 400) e V31 (Consumo de energia específico da CTMP) em que o valor encontrado para a sensibilidade das duas variáveis são negativos, isso apenas indica sentido, o que indica que ao aumentar o consumo de energia da CTMP e o torque da prensa de cavacos, mantido outras variáveis constantes, ocorrerá uma diminuição no índice de resistência à tração do produto e que, as variáveis V2 (Balança de cavacos) e V6 possuem a mesma influência sobre a resistência à tração, mas em sentidos opostos.

5. CONCLUSÕES

Através da análise do conjunto de dados foi possível observar a presença de dois padrões de comportamentos dos dados, assim identificando especificações diferentes da variável de interesse devido a necessidade da empresa em atender a demanda de qualidade dos clientes.

A análise estatística das variáveis conseguiu evidenciar variáveis no conjunto de dados na qual possuem baixa influência na resistência à tração, nos permitindo uma diminuição do grupo de variáveis, o que ajuda a poupar poder computacional na hora de desenvolver os modelos de predição.

A qualidade do conjunto de dados para construção do modelo se mostrou de grande importância pois tanto o modelo estatístico baseado em regressão linear multivariada quanto o modelo não-linear baseado em regressão por mínimos quadrados parciais não apresentaram uma capacidade satisfatória de prever a variável de interesse. O modelo baseado em redes neurais artificiais também apresentou baixo desempenho apesar de conseguir um desempenho melhor que os modelos estatísticos. Isso indica que pode haver um problema nos dados, podendo ser alguns dos problemas a calibração, falta de alguma variável importante para a predição da variável de interesse ou necessidade de uma nova abordagem sobre os dados para uma nova metodologia de tratamento de dados.

Principal Component Analysis (PCA) se mostrou um método não eficaz para a resolução do problema do conjunto de dados específico para este trabalho, visto que tanto os modelos estatísticos quanto o modelo baseado em redes neurais apresentaram resultados insatisfatórios e com capacidade de predição inferiores aos modelos previamente desenvolvidos.

A rede neural, devido ao seu alto grau de não-linearidade e de generalização se mostrou capaz de detectar os *outliers* no conjunto de dados que eram responsáveis pelo baixo desempenho dos modelos desenvolvidos. Após a remoção desses *outliers* os modelos apresentaram uma melhora significativa em seu desempenho, sendo capazes de prever com maior precisão a variável de interesse.

A análise de sensibilidade mostrou que a variável de maior sensibilidade sobre a variável de interesse é a vazão de NaOH para o silo de pré-tratamento do cavaco, indicando que a parte de tratamento químico do cavaco exerce grande influência sobre a resistência à tração.

Sugere-se como trabalhos futuros a serem executados:

- Modificar o conjunto de variáveis utilizado para o desenvolvimento deste modelo, com o objetivo de melhorar os resultados.
- Desenvolver um sensor virtual para o controle a resistência ao arrebatamento do papel cartão produzido pela Klabin, visto que é uma variável de interesse dos clientes.
- Desenvolver um modelo em outra planta da empresa no qual o processo de polpação não seja o chemi-thermomechanical pulp, para comparar o desempenho dos modelos para a predição dos valores registrados da resistência à tração.

REFERÊNCIAS

AGUIAR, F.G. **Utilização de Redes Neurais Artificiais para detecção de padrões de vazamento em dutos**. Universidade de São Paulo, São Carlos, 2010.

AGUIRRE, L.A. **Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais**. 2nd ed. Belo Horizonte: Editora UFMG, 2004.

ALMEIDA, F.S. **Construção De Uma Planta Química Piloto Para O Controle Do PH De Processos Contínuos Utilizando A Diluição Do Agente Químico De Neutralização Como Variável De Controle**. Centro Universitário Do Leste De Minas Gerais, Coronel Fabriciano, 2017.

ALMEIDA, Gustavo M.; PARK, Song W. Fault detection and diagnosis in the DAMADICS benchmark actuator system—A hidden Markov model approach. **IFAC Proceedings Volumes**, v. 41, n. 2, p. 12419-12424, 2008.

ALMEIDA, S.F. **Influência da carga alcalina no processo de polpação Lo-Solids para madeira de eucalipto**. Universidade Estadual do Norte Fluminense, Campo dos Goytacazes. Dissertação de mestrado, 2015.

ALMEIDA, R.N. **O método dos Mínimos Quadrados: Estudo e aplicações para o ensino médio**. Piracicaba-SP. Dissertação de mestrado. 2003. 115p.

ARANHA, P.S. **Análise de Correlação de Focos de Queimadas com Variáveis Climáticas no Município de Marabá**. Universidade Federal do Pará, Belém. Dissertação de mestrado. 2016.

BASSA, A.G.M.C.; SILVA JUNIOR, F.G.; SACON, V.M. Misturas De Madeira De Eucalyptus Grandis X Eucalyptus Urophylla E Pinus Taeda Para Produção De Celulose Kraft Através Do Processo Lo-Solids. **Scientia Forestalis**, v. 75, p. 19-29, 2007.

BATISTA, B. **Soluções de Equações Diferenciais Usando Redes Neurais de Múltiplas Camadas com os métodos da Descida mais íngreme e Lvenberg-Marquardt**. Universidade Federal Do Pará, Belém, 2012.

BRAGA, A.P.; LUDERMIR, T.B.; CARVALHO, A.C.P.L.F. **Redes Neurais Artificiais: Teoria E Aplicações**. LTC, Rio de Janeiro, 2000.

CARLSSON, L. A.; LINDSTROM, T. A shear-lag approach to the tensile strength of paper. **Composites Science and Technology**, v. 65, n. 2, p. 183-189, 2005.

CASELLA, G.; BERGER, R.L. **Statistical inference**. 2.ed. Pacific Grove USA: DUXBURY/Thomson Learning, 2002.

CASILLO, D. S.S. **Controle preditivo não linear baseado no modelo de Hammerstein com prova de estabilidade**. 2009. 135 f. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal do Rio Grande do Norte, Natal, 2009.

CHAKRABORTY, D. NARAYANAN, V. GHOSH, A. Integration of deep feature extraction and ensemble learning for outlier detection. **Pattern Recognition**, v. 89, p. 161-171, 2019.

CIT, E. J. **Qualidades da folha de polpa kraft em diferentes proporções de Pinus taeda L. e Eucalyptus dunnii M.** Universidade Federal do Paraná, Curitiba, 2007.

D'ALMEIDA, M.L.O. Celulose E Papel: Tecnologia De Fabricação Da Pasta Celulósica. 2.ed. **São Paulo: IPT**, 1988. p. 45-106.

DE CARVALHO, D.M.; DA SILVA, M.R.; COLODETTE, J.L. Estudo da relação entre condições de polpação e propriedades físico-mecânicas do papel. **Revista Árvore**, v. 39, n. 3, p. 575-584, 2015.

DJELLOULI, A., BENYELLOUL, K., AOURAG, H., BEKHECHI, S., ADJADJ, A., BOUHADDA, Y., & ELKEDIM, O. A datamining approach to classify, select and predict the formation enthalpy for intermetallic compound hydrides. **International Journal of Hydrogen Energy**, v. 43, n. 41, p. 19111-19120, 2018.

DUFOUR, P.; BHARTIYA, S.; DHURJATI, P.S.; DOYLE, F.J. Neural network-based software sensor: training set design and application to a continuous pulp digester. **Control Engineering Practice**, v. 13, n. 2, p. 135-143, 2005.

FACCHIN, S. **Técnicas de Análise Multivariável aplicas ao Desenvolvimento de Analisadores Virtuais**. Tese (Mestrado). Universidade Federal do Rio Grande do Sul. 2005;

FLECK, L. Redes Neurais Artificiais: Princípios Básicos. **Revista Eletrônica Científica Inovação e Tecnologia**, v. 7, n. 13, p. 47-57, 2016.

GARCIA, C. **Modelagem E Simulação De Processos Industriais E De Sistemas Eletromecânicos**. 2nd ed. São Paulo: Edusp, 2005.

GARCIA, D. **ANÁLISE DO COMPORTAMENTO DE MARMORIZAÇÃO DE CARTÃO REVESTIDO USANDO MODELAGEM MATEMÁTICA EM REDES NEURAI**S. Universidade Federal Do Paraná, Curitiba, 2010.

GELADI, L.; KOWALSKI, B. R. PARTIAL LEAST SQUARES REGRESSION: A TUTORIAL. **ANALYTICA CHIMICA ACTA** V.185, 1-17, 1986.

GOMIDE, J.L.; COLODETTE, J.L.; DE OLIVEIRA, R.C.; SILVA, C.M. Caracterização tecnológica, para produção de celulose, da nova geração de clones de Eucalyptus do Brasil. **Revista Árvore**, v. 29, n. 1, 2005.

HALKIDI, M.; BATISTAKIS, Y.; VAZIRGIANNIS, M. On clustering validation techniques. **Journal of intelligent information systems**, v. 17, n. 2-3, p. 107-145, 2001.

HAYKIN, S. **Redes neurais: princípios e prática**. ARTMED EDITORA LTDA, 2nd ed. Porto Alegre, 2001.

HONGYU, K.; SANDANIELO, V. L. M.; DE OLIVEIRA JUNIOR, G. J. Análise de componentes principais: resumo teórico, aplicação e interpretação. **E&S Engineering and Science**, v. 5, n. 1, p. 83-90, 2016.

HUSSAIN, M. A. Review of the applications of neural networks in chemical process control—simulation and online implementation. **Artificial intelligence in engineering**, v. 13, n. 1, p. 55-68, 1999.

JAIN, A. K. Data clustering: 50 years beyond K-means. **Pattern Recognition Letters**, n. 31, p. 651-666, 2010.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. Madison: Prentice Hall International, 1998. 816p.

JOLLIFFE, I.T. Discarding Variables in a Principal Component Analysis. II: Real Data. **Journal of the Royal Statistical Society**, v. 22, n. 1, p. 21–31, 1973.

KAISER, H. F. The Varimax criteria for analytical rotation in factor analysis. **Psychometrika**, v. 23, n. 3 p. 141-151, 1958.

KANEKO, H; ARAKAWA, M; FUNATSU, K. Development of a new soft sensor method using independent component analysis and partial least squares. **AIChE Journal**, v. 55, n. 1, p. 87-98, 2009.

KLEN, A. M.; LANA, M. S. Fuzzy Algorithm of discontinuity sets. **Rem: Revista Escola de Minas**, v. 67, n. 4, p. 439-445, 2014.

KLOCK, U.; ANDRADE, A.; HERNANDES, J. **Propriedades do Papel**. Série Didática; FUPEF: Curitiba, Brazil, 2013;

KUMAR, A.; ABIRAMI, S. Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method. **Information Sciences**, v. 460, p. 23-41, 2018.

Larsson, P. T., Lindström, T., Carlsson, L. A., Fellers, C. Fiber length and bonding effects on tensile strength and toughness of kraft paper. **Journal of materials science**, v. 53, n. 4, p. 3006-3015, 2018

LIU J.L. On-line soft sensor for polyethylene process with multiple production grades. **Control Eng Practice**. 2007; 15:769–778.

LIU, Z.; GE, Z.; CHEN, G.; SONG, Z. Adaptive soft sensors for quality prediction under the framework of Bayesian network. **Control Engineering Practice**, v.72, p. 19-28, 2018.

LOPEZ-HIDALGO, A. M.; ALVARADO-CUEVAS, Z. D.; DE LEON-RODRIGUEZ, A. Biohydrogen production from mixtures of agro-industrial wastes: Chemometric analysis, optimization and scaling up. **Energy**, v. 159, p. 32-41, 2018.

LOTUFO, F.A.; GARCIA, C. Sensores Virtuais ou Soft Sensors: Uma Introdução. In: **Proceedings of the 7th Brazilian Conference on Dynamics, Control and Applications**, Presidente Prudente. 2008. p. 1-9.

MATIAS, T.; SOUZA, F.; ARAÚJO, R.; ANTUNES, C. H. Learning of a single-hidden layer feedforward neural network using an optimized extreme learning machine. **Neurocomputing**, v. 129, p. 428-436, 2014.

METZ, J. **Interpretação de clusters gerados por algoritmos de clustering hierárquico**. 2006. Tese de Doutorado. Universidade de São Paulo.

MORELLATO, S. A. **Modelos de Regressão PLS com Erros Heteroscedásticos**. Universidade Federal de São Carlos, São Carlos. Dissertação de mestrado, 2010.

PARMAR, J.D.; PATEL, J.T. Anomaly Detection in Data Mining: A Review. **International Journal of Advanced Research in Computer Science and Software Engineering**, v. 7, n. 4, 2017.

POPPI, R.J.; DE ANDRADE, J.C.; DE CERQUEIRA, E.O. Redes Neurais E Suas Aplicações Em Calibração Multivariada. **Química Nova**, V. 24, n. 6, p. 864-873, 2001.

PUTH, M.; NEUHÄUSER, M.; RUXTON, G. D. Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits. **Animal Behaviour**, v. 102, p. 77-84, 2015.

RIBEIRO, R. N.; ALBUQUERQUE, R. Utilização de Redes Neurais Artificiais para Predição do Grau de Cozimento de um Digestor de Celulose. **SEGeT–Simpósio de Excelência em Gestão e Tecnologia**, 2007.

ROTARU, P.C. Empirical study on regional employment rate in Romania. **Procedia-Social and Behavioral Sciences**, v. 109, p. 1365-1369, 2014.

SAMISTRARO, G.; MUNIZ, G. I. B. Previsão Das Propriedades Físicas Do Papel Kraft Por Espectroscopia No Infravermelho Próximo (NIR) E Regressão Por Mínimos Quadrados Parciais (PLS). **Química Nova**, v. 32, p. 1422-1425, 2009.

SCOTT, W.; SHIRT, R. Potential application of predictive tensile strength models in paper manufacture: Part II – Integration of a tensile strength model with a dynamic paper machine material balance simulation, **TAPPI Papermakers Conference**, p. 879–887, 2001.

SHARIFI, T. H.; HAMZEH, Y.; POURMOUSA, S. H. Optimization Process Variables Of Deinked Pulp Bleaching-Improvement Of Wet Tensile Strength Of Tissue Paper. **Iranian Journal of Wood and Paper Science Research**, v. 30, 2016.

SILVEIRA, F.G. **Abordagem Geométrica do Método dos Quadrados Mínimos Parciais com uma Aplicação a Dados de Seleção Genômica**. Universidade Federal De Lavras, Lavras, 2014.

SMOOK, G. A. **Handbook for pulp & paper technologists**. 2nd. Angus Wilde Publications Inc. Vancouver, 1992. 419p.

SOARES, F. M.; OLIVEIRA, R. C.L. Inferência de temperatura de Fornos de Redução de Alumínio Primário através de Sensores Virtuais. In: **IX Congresso Brasileiro de Redes Neurais**, Ouro Preto-MG. 2010. p. 25-28.

STEIN, F.R. **Modelagem de produção industrial de celulose Kraft com modelos aditivos generalizados e Redes Neurais**. Universidade Federal De Viçosa, Viçosa, 2010.

UPASANI, N; Om, H. Evolving fuzzy min-max neural network for outlier detection. **Procedia computer science**, v. 45, p. 753-761, 2015.

VIEIRA, O. **Construção de modelos empíricos e projeto da operação de uma máquina de produção de cartão embalagem**. Tese (Doutorado). Universidade Federal do Rio de Janeiro. 2003;

WANG, H.; JIE, G.; SHANSHAN, W.; SAPORTA, G. Spatial partial least squares autoregression: Algorithm and applications. **Chemometrics and Intelligent Laboratory Systems**, v. 184, p. 123-131, 2019.

WOLD, Herman. Path models with latent variables: The NIPALS approach. In: **Quantitative sociology**. Academic Press, 1975. p. 307-357.

XUEFENG, Y. Hybrid artificial neural network based on BP-PLSR and its application in development of soft sensors. **Chemometrics and Intelligent Laboratory Systems**, v. 103, n. 2, p. 152-159, 2010.

ZANATA, D.R.P. **Desenvolvimento de sensor virtual empregando redes neurais para medição da composição em uma coluna de destilação**. 2005. Tese de Doutorado. Universidade de São Paulo.

ZANUNCIO, A.J.V.; COLODETTE, J.L.; DE OLIVEIRA, R.C. Características químicas da polpa celulósica de eucalipto e as propriedades ópticas e físico-mecânicas do papel. **Revista Ciência da Madeira (Brazilian Journal of Wood Science)**, v. 2, n. 2, p. 10-12953/2177-6830. v02n02a02, 2011.

ZHU, S; HONGGUI, H; MIN, H; JUFEI, Q. A data-derived soft-sensor method for monitoring effluent total phosphorus. **Chinese journal of chemical engineering**, v. 25, n. 12, p. 1791-1797, 2017.

ZYNGIER, D. **Desenvolvimento e implementação de sensores virtuais para monitoramento de um bioprocesso**. 2000. Tese de Doutorado. Universidade Federal do Rio de Janeiro.